Snap DSA Report

Risk Assessment Results and Mitigations



25 August 2026

Foreword **5**
What's New? **6**
What is New in this Report? 6
1. Introduction **8**
1.1 Snapchat 101 8
1.2 Snapchat Community 9
2. DSA Risk Assessment Scope **9**
2.1 Approach 9
2.2 Spotlight 9
2.3 Discover 11
2.4 Public Profiles 13
2.5 Snap Map 15
2.6 Lenses 17
2.7 Advertising 20
3. DSA Risk Assessment Methodology **23**
3.1 Identification of Risks 24
3.2 Likelihood Analysis 25
3.3 Severity Analysis 27
3.4 DSA Risk Factors 27
3.5 Overall Potential Risk Prioritization 28
3.6 Snap's Mitigations 29
3.7 Conclusions 29
4. DSA Risk Assessment Results **29**
4.1 Category 1 - Dissemination of Illegal Content or Content that Violates Snap's Terms and Conditions 30
4.1.1 Dissemination of Child Sexual Abuse Material 33
4.1.2 Dissemination of Illegal Hate Speech 42
4.1.3 Dissemination of Information Related to the Sale of Prohibited Products or Services 50
4.1.4 Dissemination of Terrorist Content 63
4.1.5 Dissemination of Content that Infringes on Intellectual Property Rights 71
4.1.6 Dissemination of Adult Sexual Content 78
4.1.7 Dissemination of Content regarding Harassment and Bullying 88
4.1.8 Dissemination of Content that Glorifies Self-Harm, including the Promotion of Self-Injury, Suicide or Eating Disorders 96
4.1.9 Dissemination of Content encouraging or engaging in Violent or Dangerous Behaviour 105
4.1.10 Dissemination of Harmful False Information 115
4.1.11 Dissemination of Fraud and Spam 123
4.1.12 Dissemination of Information Related to Other Illegal Activities 134

4.2 Category 2 - Negative Effects on Fundamental EU Rights....141
4.2.1 Right to Human Dignity....144
4.2.2 Right to Freedom of Expression and Assembly....148
4.2.3 Right to Private Life....156
4.2.4 Right to the Protection of Personal Data....164
4.2.5 Right to Non-Discrimination and Freedom of Religion....173
4.2.6 Children's Rights....182
4.2.7 Right to Consumer Protection....191
4.2.8 Right to Property....200
4.3 Category 3 — Negative Effects on Democratic and Electoral Processes, Civic Discourse, and Public Security....202
4.3.1 Negative Effects on Democratic and Electoral Processes....203
4.3.2 Negative Effects on Civic Discourse....209
4.3.3 Negative Effects on Public Security....218
4.4 Category 4 - Negative Effects on Public Health....227
4.4.1 Negative Effects on Public Health....228
4.4.2 Negative Effects on Gender-Based Violence....234
4.4.3 Negative Effects on Minors....244
4.4.4 Serious Negative Consequences on Physical and Mental Well-Being....267
5. Specific Mitigations....281
5.1 Snapchat Design and Function....283
5.1.1 Introduction....283
5.1.2 Oversight and Administration....283
5.1.3 Adaptations and Mitigations....284
5.1.4 Integrations with other mitigations....289
5.1.5 Online Interface Design Process....290
5.1.6 Conclusion....290
5.2 Terms....290
5.2.1 Introduction....290
5.2.2 Oversight and Administration....292
5.2.3 Terms and Conditions....292
5.2.4 Accessing Terms and Conditions....296
5.2.5 Languages....297
5.2.6 Readability....297
5.2.7 Conclusion....298
5.3 Transparency....298
5.3.1 Information provided on the website....298
5.3.2 Information provided in app stores....300
5.3.3 Information provided in the application....301

5.3.4 Languages..... 303
5.3.5 Conclusion..... 303
5.4 Content Moderation..... 304
5.4.1 Approach..... 304
5.4.2 Oversight and Administration..... 304
5.4.3 Proactive Moderation..... 306
5.4.4 Search and contact surface moderation..... 309
5.4.5 Product-Specific Moderation..... 311
5.4.6 Reactive Moderation (Reporting)..... 312
5.4.7 Moderation Effectiveness Metrics..... 315
5.4.8 Conclusion..... 316
5.5 Enforcement..... 316
5.6 Algorithmic Systems..... 337
5.7 Advertising Systems..... 353
5.8 Protection of Minors..... 370
5.9 Content Authenticity..... 393
5.10 Trusted Flaggers..... 404
5.11 Dispute Settlement Bodies..... 414
5.12 Codes and Crisis Protocols..... 418
6. Ongoing Risk Detection and Management..... 427
6.1 Platform Principles-based Framework..... 427
6.2 DSA Compliance Function and Cross-Functional Working Groups..... 432
6.3 Privacy and Safety by Design..... 442
6.4 Prevalence Testing..... 449
6.5 External Request Monitoring and Review..... 454
6.6 Digital Well-Being Index (DWBI) Initiative..... 459
6.7 Snap Advisory Groups..... 464
6.8 Audit..... 473
Additional Independent Assurance..... 474
7. Conclusion..... 475
8. Final Words..... 478

Foreword

Snap's purpose is to empower people to express themselves, live in the moment, learn about the world, and have fun together. Our flagship product, Snapchat, is built around the camera - a fundamentally different design philosophy from traditional social media. It opens to a camera, not a public feed. Conversations delete by default. Its primary use is private messaging between close friends and family. That design reflects a deliberate commitment: to give people a platform where they can communicate freely without the pressures of a public performance. Safety and privacy-by-design are central to Snapchat.

This is Snap's Fourth Annual Risk Assessment and Mitigation Report, submitted pursuant to Article 42(4) of Regulation (EU) 2022/2065 (the Digital Services Act). Snapchat was designated a Very Large Online Platform by the European Commission on 25 April 2023, with obligations taking effect in September 2023, on the basis that the average number of monthly active recipients of Snapchat in the European Union exceeds 45 million. This Report covers the annual assessment period ending 30 June 2026 and is prepared against the baseline of the Third Annual Report (25 August 2025).

The designation is limited to services provided as part of Snapchat that meet the definition of "online platform" in Article 3, point (i) of Regulation (EU) 2022/2065 (the Digital Services Act, DSA): hosting services that store and disseminate information to the public at the request of a recipient. As recital 14 of the DSA and the designation decision make explicit, interpersonal communication services fall outside that definition because they are used for communication between a finite number of persons determined by the sender, not dissemination to the public at large. There has been no change to Snapchat's designation status since the Third Risk Assessment.

Nor has there been a change to the scope of Snap's risk assessment exercise since the Third Risk Assessment. For the purposes of this risk assessment in relation to the Snapchat service in its current form, the six features which Snap considers are Spotlight, Discover, Public Profiles, Snap Map, Lenses, and Advertising. When referring to "in scope" services, surfaces or functionalities in this Report, we are referring to the scope of Snap's risk assessment exercise, rather than taking a formal legal position about the scope of the "online platform" definition in the DSA.[1] Given Snap's broader and holistic commitment to safety, our risk assessment and mitigation exercise may extend to surfaces or functionalities which are wholly or partially outside the six features listed above.

The 'What's New?' section summarises the principal changes in this Report relative to the Third Risk Assessment. Section 1 provides an updated introduction to the Snapchat platform and community. Section 2 sets out the scope of the risk assessment. Section 3 describes the

[1] Unless indicated otherwise, the terms service, surface and functionality are used interchangeably in this report, without intending to indicate a specific legal delineation.

[redacted]

methodology. Section 4 contains the risk assessment results, including the new or expanded risks identified through this assessment cycle. Section 5 sets out Snap's specific mitigations. Section 6 describes Snap's ongoing risk-detection and governance framework.

We are pleased to note that, since our 2025 Risk Assessment, our overall Policy Violating Prevalence rate-the primary metric we use to assess the prevalence of violating content-declined [redacted] This reduction is consistent with continued improvements to detection mechanisms, reviewer training and enforcement processes across the public-content surfaces included in Snap's prevalence testing.

Snap's second external DSA audit concluded that Snap complied with the applicable DSA specified requirements in all material respects. During the 2026 assessment period, Snap also introduced or strengthened a number of product safeguards, particularly for minors. These included more privacy- and safety-protective default settings and a redesigned onboarding flow for minor accounts; refinements to friending protections and warning notices for higher-risk contact; improvements to the prominence and usability of the illegal-content reporting pathway; a redesigned Heads Up intervention for drug-related searches and profiles; and additional steps to reduce sensitive content shown to users under 18 on Spotlight. [redacted] Wholly AI-generated videos are no longer eligible for recommendation on Spotlight, while content enhanced or edited using Snapchat's AI creative tools remains eligible for recommendation and includes transparency indicators. This change was implemented after the assessment cutoff and is therefore noted as a subsequent development rather than a mitigation in place during the 2026 assessment period. These changes are described in further detail in Section 5 of this Report.

Snap looks forward to continuing its constructive dialogue with the Commission, our Digital Services Coordinator, and other regulators. We remain committed to the objectives of the DSA and to demonstrating, through this Report, that Snapchat's approach to systemic risk is rigorous, proportionate, and responsive to regulatory feedback.

What's New?

What is New in this Report?

This Report reflects the outputs of Snap's annual DSA systemic risk assessment cycle and is prepared against the baseline of Snap's Third Annual Risk Assessment and Mitigation Report dated 25 August 2025 (the "Third Risk Assessment"). The most material updates to this Report are concentrated in Sections 4 and 5.

Section 1 - Introduction

[redacted]

In all material respects, Snapchat's services assessed in this report remain as described in the Third Risk Assessment. [redacted] While not considered likely to have a critical impact on the risks identified pursuant to Article 34 of the DSA, a new feature that Snap began rolling out in June 2026 is a Friend-Only content profile for users aged 13–15 to create, save and showcase Stories and Spotlight videos on a dedicated profile visible only to their mutually accepted friends, without public engagement metrics such as favorite counts. Following the assessment cutoff, the rollout progressed such that, as of July 2026, users aged 13–15 could no longer post publicly to Spotlight.

Community statistics have been updated to reflect current EU monthly active recipient figures and age and gender distribution; [redacted]

Section 2 - DSA Risk Assessment Scope

Snapchat's designation status has not changed since September 2023. The Commission has not issued guidance since the Third Risk Assessment that would alter Snap's scope analysis. Snap continues to assess Spotlight, Discover, Public Profiles, Snap Map, Lenses, and Advertising in the risk assessment and mitigation exercise.

Section 3 - DSA Risk Assessment Methodology

The risk assessment methodology builds on that of the Third Risk Assessment, with several refinements. The methodology continues to assess likelihood, severity and overall potential risk prioritisation across the four DSA risk categories, with explicit treatment of the Article 34(2) risk factors for each harm.

Section 4 - DSA Risk Assessment Results

This section contains the most significant updates relative to the Third Risk Assessment, arising from Snap's updated assessment across the relevant risk categories. Compared to the Third Risk Assessment, this Risk Assessment provides more details on potential under-13 access, which incorporates publicly available survey evidence concerning Snapchat access and use by children below or around the minimum-age threshold, as well as an expanded analysis of the risk of users misrepresenting their age. This Risk Assessment also includes an assessment of non-coercive criminal recruitment of minors, analysed separately from the existing human-trafficking risk.

In addition to these new and expanded considerations, likelihood, severity, and overall risk prioritisation assessments have been reviewed and updated across all risk categories.

Section 5 - Specific Mitigations

[redacted]

Find Friends moderation mechanisms, the text moderation system, broadcast content moderation, and the Heads Up tool have been documented with greater specificity than in prior Risk Assessments, in order to provide a fuller evidentiary basis for proportionality assessment. Other mitigations remain substantively as described in the Third Risk Assessment, with updates reflecting product and operational changes implemented during the 2026 assessment period. In particular, Section 5 records several enhancements to the protection of minors, including refinements to minor-account default settings and the onboarding flow; continued improvements to Family Center accessibility and discoverability, including additional in-app entry points; friending-protection refinements; and expanded profile-level moderation across usernames, display names and profiles. The Section also records improvements to reporting and notice-and-action pathways, together with additional mitigation updates, [redacted] and the extension of native reporting functionality to web.snapchat.com.

Separately, Section 5 reflects a material change to public-content protections for younger teens. In June 2026, Snap began rolling out a Friend-Only content profile for users aged 13–15, allowing Stories and Spotlight videos to be shared on a dedicated profile visible only to mutually accepted friends. The rollout continued after the assessment cutoff and, as of July 2026, users aged 13–15 could no longer post publicly to Spotlight.

Section 6 - Ongoing Risk Detection and Governance

There have been no material changes to Snap's governance framework. The Cross-Functional DSA Governance Team continues to operate as described in the Third Risk Assessment. Snap's Safety Operations Outreach function conducted an engagement tour across Norway, Sweden, and Denmark in April 2026, establishing direct operational contact with SPOC units, KRIPOS, Europol OTF GRIMM, the Swedish IRU, Säpo, and PET. Nordic law enforcement stakeholders reported high satisfaction with Snap's operational responsiveness. Direct channels were established with Europol OTF GRIMM and the Swedish IRU in relation to Violence-as-a-Service recruitment ecosystems involving minors.

1. Introduction

1.1 Snapchat 101

Snapchat is a visual messaging and social media application developed and operated by Snap Inc. A description of the Snapchat platform and its services is available on Snap's EU Transparency page and its public support pages[2]. Snap's previous Annual Risk Assessments

[2] https://values.snap.com/privacy/transparency

provide further detail on each of the in-scope services relevant to this Report; this Report does not reproduce that description. The scope of services covered by this Risk Assessment is set out in Section 2.

1.2 Snapchat Community

EU average monthly active recipient figures are published on Snap's EU Transparency page pursuant to Article 24(2) DSA and are updated every six months. This Report refers to those published figures rather than reproducing them in full.

For the most recent measurement period, users aged 13–17 represented approximately [redacted] of Snapchat's EU user base, compared with approximately [redacted] in prior reporting periods. This demographic composition is a relevant input to the likelihood analysis in Section 4, including the assessment of risks affecting minors.

2. DSA Risk Assessment Scope

2.1 Approach

The scope of this Risk Assessment is based on the Commission's designation decision of 25 April 2023 and the definition of 'online platform' in Article 3, point (i) DSA. That scope has not changed since the Third Risk Assessment. Since the Third Risk Assessment: (i) Snapchat's designation status has not changed; (ii) the Commission has not issued guidance that would materially alter Snap's scope analysis; and (iii) the functionality of Snapchat's in-scope features has not changed in a way that causes Snap to change the scope of the risk assessment process.

Snap continues to assess the following six features to fall in the risk assessment and mitigation exercise. Brief descriptions of each feature, and any changes since the Third Risk Assessment, are set out in the subsections below.

2.2 Spotlight

What is Spotlight?

Spotlight is the public, short-form video surface within Snapchat. Spotlight was launched in November 2020. Users can submit videos to Spotlight from within the Snapchat app or via web. Videos submitted to Spotlight are public: they are accessible to other users within the Snapchat app, on the web, and a link to a Spotlight video can be shared to other platforms.

How does Spotlight work?

Spotlight presents users with a personalised feed of short-form videos. The ranking applied to the feed is described in the public ranking-methodology article "How We Rank Content on

[redacted]

Spotlight," referenced in the Terms of Service. Users can opt out of personalisation through the Snapchat settings; the personalised content recommendation may then be replaced by a non-personalised content recommendation, as further described in Section 5.6 (Algorithmic Systems).

Content submitted to Spotlight is subject to the moderation framework described in Section 5.4 (Content Moderation). [redacted]

Functionality

Spotlight functionality, as a general matter, includes the following:

- comments may be added to Spotlight videos. Comments [redacted] may then accept, reject or auto-approve comments. Once accepted or auto-approved, the comment is publicly visible on the Spotlight video. Comments can be deleted or reported by the creator and reported by viewers;
- viewers can indicate fondness for a Spotlight video using the heart ("favorite") icon;
- a Spotlight video can be reported by a viewer using the in-app reporting menu, including under the illegal content reporting option (further described in Section 5.4.6);
- a creator can delete their Spotlight submissions individually;
- users have the ability to share Spotlight content to their Story
- Users can also repost a Spotlight video to share it with others. Reposts are public and can be seen on the users Public Profile or in the Spotlight feed.
- [redacted].
- content on Spotlight does not auto-advance, meaning users are required to actively tap to go to the next piece of Spotlight content. If the user takes no action, the same piece of content loops.

Spotlight Monetization Program

The Spotlight Rewards Program ended on 31 January 2025. Eligible creators invited to participate in Snap's unified Monetization Program may continue to receive rewards in connection with

[redacted]

qualifying Spotlight content, subject to the applicable Monetization Terms. Spotlight content remains subject to Snap's Terms of Service, Community Guidelines, and Content Guidelines for Recommendation Eligibility.[3]

Snapchat+ features relevant to Spotlight

As of the 2025 Report, no Snapchat+ features were specific to Spotlight.

Generative AI features relevant to Spotlight

Spotlight does not, as a general matter, include a generative-AI content creation feature specific to the Spotlight surface. Certain generative-AI integrations may nevertheless interact with Spotlight content, including the ability to tag My AI in Spotlight comments and to share certain Spotlight videos with My AI. In addition, content created using Snap generative-AI tools may be submitted to and disseminated through Spotlight. The associated risk implications and mitigation measures are addressed in the relevant Section 4 assessments and in Section 5.9 (Content Authenticity).

[redacted][4]

2.3 Discover

What is Discover?

Discover is a public content surface within the Stories tab of the Snapchat app. It presents content from creators and Media Partners. The product has at times been referred to as "For You"; any reference to "For You" in this Report or in prior Reports should be read as a reference to Discover.

Discover comprises two principal content streams: (i) content from Media Partners onboarded to Snapchat through the Partnerships team; and (ii) user-generated content (UGC) from Snap Stars and from Popular Users (Snapchatters whose Story settings are public and who meet a

[3] https://help.snapchat.com/hc/en-us/articles/7012310905876-How-does-Snapchat-determine-who-is-eligible-to-participate-in-the-Spotlight-Reward-Program

[4] https://newsroom.snap.com/rewarding-authentic-creativity-on-spotlight

[redacted]

follower-count threshold). Discover content is, as a general matter, accessible to users including those aged 13–17, subject to the audience- and ranking-level restrictions described in Section 5.6 (Algorithmic Systems).

How does Discover work?

Discover presents users with a personalised feed of creator and Media Partner content. The ranking applied to the feed is described in the public ranking-methodology article "How We Rank Content on Discover,"[5] referenced in the Terms of Service. Users can opt out of personalisation through the Snapchat settings; the personalised content recommendation may then be replaced by a non-personalised content recommendation, as further described in Section 5.6.

Content presented on Discover is subject to the moderation framework described in Section 5.4 (Content Moderation). In summary:

- **Media Partner content.** Media Partners are onboarded by the Partnerships team prior to being permitted to distribute content on Discover. [redacted]
- **User-generated content (Snap Stars and Popular Users).** UGC submitted by Snap Stars and Popular Users is subject to automated moderation and, where relevant, human moderation prior to becoming eligible for Discover. [redacted]

[redacted]

Functionality

Discover functionality, as a general matter, includes the following:

- viewers may send replies to eligible Public Stories. Story Replies are visible to the creator and are not publicly displayed unless the creator elects to quote or share a reply in their

[5] https://help.snapchat.com/hc/en-us/articles/8961631424020-How-We-Rank-Content-on-Discover?

[redacted]

Public Story. Story Replies are subject to automated spam and abuse controls. Comments may be added to Publisher videos in Discover. Comments [redacted], who may then accept, reject or auto-approve comments. Once accepted or auto-approved, the comment is publicly visible. Comments can be deleted or reported by the creator and reported by viewers.

- view counts are not displayed on Publisher Stories or Public Stories;
- publication duration for Stories is limited. Snapchat+ subscribers may have access to functionality that permits them to adjust when Story Snaps expire;
- Discover content does not automatically advance from one user's content to the next user's content;
- a user's likes ("favorites") are not publicly displayed;
- Discover content can be reported by a viewer using the in-app reporting menu, including under the illegal content reporting option (further described in Section 5.4.6).

Snapchat+ features relevant to Discover

As of the 2025 Report, no Snapchat+ features were specific to Discover.

Generative AI features relevant to Discover

Discover does not, as a general matter, include a generative-AI content creation feature specific to the Discover surface. Content created using Snap generative-AI tools may nevertheless be disseminated through Discover. The associated recommendation-eligibility considerations, risk implications, and mitigation measures are addressed in the relevant Section 4 assessments and in Section 5.9 (Content Authenticity).[6]

2.4 Public Profiles

What are Public Profiles?

Public Profiles enable eligible Snapchatters aged 16 and older in certain countries to share content publicly and grow an audience on Snapchat. All Snapchatters have a private profile by default, and use of the Public Profile and the decision to post content publicly remain optional. A Public Profile can display profile information, Saved Stories, Spotlight content, reposts and Lenses. [redacted]

How do Public Profiles work?

Content displayed on a Public Profile is subject to the moderation framework described in Section 5.4 (Content Moderation). [redacted]

[6] https://values.snap.com/policy/content-guidelines-recommendation-eligibility/recommendation-eligibility/quality?.

[7] https://help.snapchat.com/hc/en-us/articles/7012292808596-About-Profiles-on-Snapchat.

[redacted]

[redacted]

Eligible users can access their Public Profile through the My Profile screen or by attempting to post publicly for the first time, and can choose whether to populate it with profile information or public content. Users retain control over whether and where individual Snaps are shared and whether content is saved to their Public Profile. [redacted]

Functionality

Public Profile functionality, as a general matter, includes the following:

- a user can populate the Public Profile with a display name, username, bio, description, location and profile photo;
- a user can save Spotlight content, Saved Stories and Lenses to the Public Profile, and can choose to display reposts;
- a user can be followed by other Snapchatters; the follower count may be displayed on the Public Profile;
- a user can access Public Story Insights for content associated with the Public Profile;
- a user can clear their Public Profile in a single action, permanently removing information and content added to the Public Profile, including the bio, profile photo and saved Public Stories. Clearing a Public Profile does not delete the user's Snapchat account, username, Spotlight posts or Snap Map posts, which must be managed separately;
- a user can disable remixes of their public content;
- a user can disable, filter or block Story replies, and can manage who may contact them through privacy settings;
- [redacted];
- a Public Profile (including content displayed on it and the account itself) can be reported by a viewer using the in-app reporting menu, including under the illegal content reporting option (further described in Section 5.4.6).

Account upgrades

Some eligible Public Profile users may separately apply for verification as a Snap Star. These account types provide additional tools intended to support audience growth and engagement. Snap Star eligibility is assessed separately by reference to applicable program criteria.

[8] https://help.snapchat.com/hc/en-us/articles/14855534914708-Your-Privacy-When-You-Post-Publicly-on-Snapchat?.

[redacted]

Snapchat+ features relevant to Public Profiles

As of the 2025 Report, no Snapchat+ features were specific to Public Profiles.

Generative AI features relevant to Public Profiles

Public Profiles do not, as a general matter, incorporate generative-AI features within the surface itself. Snap recognises that content created using generative-AI tools on Snapchat may be displayed via a Public Profile. The risk implications of generative-AI content disseminated via in-scope services, and the corresponding mitigation measures, are addressed in the relevant Section 4 assessments and in Section 5.9 (Content Authenticity).

2.5 Snap Map

What is Snap Map?

Snap Map is an interactive map surface within Snapchat. It surfaces user-generated content tagged to locations and, where the user has opted in, the locations of bidirectional friends. Snap Map is available to users aged 13 and over.

Location sharing on Snap Map is between bidirectional friends. There is no functionality to share location with non-friends. The location-sharing dimension of Snap Map (as a friend-to-friend feature) is, in itself, out of scope of this Report; the user-generated public content surfaces of Snap Map (Map Stories, Place Stories, City Stories, Heat Map, and tagged Friend Stories) are in scope. The friend-to-friend location-sharing safeguards are nonetheless described in Sections 5.1 and 5.3 because they affect the overall Snap Map user experience and bear on minor-protection mitigations.

How does Snap Map work?

Location sharing is, by default, off for all users ("Ghost Mode"). A user does not appear on a friend's Snap Map until they open the application and affirmatively opt in to location sharing. A user may, at any time, update their location-sharing settings, including by adjusting whether background or only-while-using-app location is shared, and which friends can see their location.

Public content is shared to Snap Map by selecting "Snap Map" on the "Send To" page. A user with a Public Profile, or a user sharing My Story to Everyone, may also have their Snaps shared to Snap Map when tagged to a place or venue, subject to the minor-account restrictions described in Section 2.4 (Public Profiles) and in Section 5.1.3.[9]

Content presented on Snap Map is subject to the moderation framework described in Section 5.4 (Content Moderation). Content posted to Snap Map is, in the ordinary course, scanned

[9] https://help.snapchat.com/hc/en-us/articles/7012271195796-How-do-I-submit-a-Snap-to-Snap-Map?.

[redacted]

asynchronously and proactively against automated detection databases, [redacted] Place Stories and similar content tagged to specific business or geographic locations are subject to additional human moderation as a condition of broader visibility.

Snap Map ranks public content using factors that vary by content type, including recency, engagement, whether the content was posted by a friend or followed creator, content category, and content metadata. The ranking applied to Snap Map is described in the public ranking-methodology article "How We Rank Content on Maps," referenced in the Terms of Service. Public content shown on Snap Map is not, in the ordinary course, personalised based on inferred user interests[10].

Content surfaces on Snap Map

Snap Map surfaces include the following user-generated content streams:

- **Map Stories.** Thumbnails on the map highlighting events or popular places.
- **Place Stories.** Stories anchored to specific places or venues, surfaced on the Map and in the corresponding Place Profiles. Place Stories may include Spotlight, Snap Map, and Public Story Snaps tagged with a place; Snaps may appear in a Place Story for up to 90 days.
- **City Stories.** Aggregations of Snaps from a locality (city, neighbourhood, state or country) over the past seven days, presented at the header of the Map.
- **Heat Map.** Visual indicators of areas with recent, high-volume content shared to Public Stories. Content up to seven days old is reflected on the Heat Map.
- **Friend Stories tagged with Places.** A view of Snaps that have been tagged to a place by a user's bidirectional friend, displayed alongside the friend's Bitmoji.

Functionality

Snap Map functionality, as a general matter, includes the following:

- comments are not, as a general matter, available on Snap Map content;
- view counts are not displayed on Snap Map Stories;
- a user can remove a Snap they submitted to Snap Map (or place-tagged in Spotlight) at any time via their profile;
- Snap Map content can be reported by a viewer using the in-app reporting menu, including under the illegal content reporting option (further described in Section 5.4.6);
- [redacted]

[10] http://help.snapchat.com/hc/en-us/articles/35157584066964-How-We-Rank-Content-on-Maps.

[redacted]

- Public Profile and My Story posting eligibility for users aged 16–17 is more limited; under-16 accounts cannot have a Public Profile and Public Stories and Spotlights from under-16 accounts are not featured on Snap Map when tagging a place or venue.

[redacted]

Snapchat+ features relevant to Snap Map

Snapchat+ subscribers have access to the following Snap Map features:

- **3D Homes.** Subscribers can select and customise a home on Snap Map. The home may be visible to friends with whom the subscriber shares location, including when the subscriber is not at that location;[11]
- **Footsteps.** Subscribers can review the places they have explored. Footsteps are visible only to the user and may be disabled or cleared; and
- **Bitmoji customisations.** Certain pet and car options are available in connection with a user's Bitmoji on Snap Map. Some pets and all currently available cars are limited to Snapchat+ subscribers.

Generative AI features relevant to Snap Map

Snap Map does not, as a general matter, include a generative-AI content-creation feature specific to the Map surface. Content created using Snap generative-AI tools may nevertheless be submitted to and disseminated through Snap Map's public-content surfaces. The associated risk implications and mitigation measures are addressed in the relevant Section 4 assessments and in Section 5.9 (Content Authenticity).

2.6 Lenses

What are Lenses?

Lenses are augmented reality (AR) experiences delivered through the Snapchat camera. Lenses can apply 3D effects, digital objects, characters and image- or voice-based transformations to a Snap. Lenses are accessed by users through the Lens Carousel, the Search interface and the

[11] https://help.snapchat.com/hc/en-us/articles/29443155505428-What-are-3D-Homes-on-Snap-Map?.

[redacted]

Lens Explorer. Sponsored Lenses are also made available to advertisers as an advertising format; the advertising-format treatment of Sponsored Lenses is described in Section 2.7 (Advertising) and in Section 5.7 (Advertising Systems).

Lenses are created either by community developers using Lens Studio (Snap's public AR development platform) or by Snap's internal Lens Team using a combination of Lens Studio and internal-only AR and generative-AI tools. Lenses are created by community developers and by Snap using Lens Studio, Snap's public AR development platform. Lens Studio includes conventional AR development tools as well as AI-assisted and generative-AI creation capabilities. Snap's internal Lens Team may also use internal AR and generative-AI tools. Lens submissions are subject to applicable submission-review and moderation processes, and new internal AR and AI capabilities are subject to applicable privacy and legal review.

How do Lenses work?

Lenses generally rely on object detection rather than facial identification. Object detection enables a Lens to recognise that a face, or a feature within a face (such as a nose or eye), is present in a frame, but does not identify or verify the identity of any specific individual. [redacted] AR development tools used in Lenses are reviewed by privacy engineering and legal teams prior to use.

Lens content is subject to the moderation framework described in Section 5.4 (Content Moderation). [redacted]

Lens types

Lenses include the following principal types:

- **Face Lenses.** Use facial-landmark tracking to position digital effects. Face Lenses do not identify or verify an individual's identity.
- **World Lenses and Landmarker Lenses.** Use environmental mapping, surface detection and spatial tracking to anchor virtual objects in the user's surroundings.
- **Scan Lenses.** Analyse objects within the camera frame to trigger specific AR experiences.
- **Connected Lenses.** Enable shared, real-time AR sessions between bidirectional friends; [redacted].

[redacted]

Lens Studio and the developer ecosystem

Lens Studio is Snap's public AR development platform. Lens Studio is available as a desktop application, with mobile and web variants. Independent developers and Snap's internal Lens Team use Lens Studio to create and publish Lenses. Lens Studio integrates a number of internal AR/ML capabilities [redacted]

Submissions to Lens Studio are subject to the Lens Studio Terms (effective 16 September 2024) and the Lens Studio Submission Guidelines. Submissions involving generative-AI tooling are additionally subject to the Generative Lens Tools Terms.

Privacy and data-handling features specific to Lenses

The following data-handling features apply to Lenses, as a general matter:

- voice-data collection for users in the European Union is, by default, off and is processed only to provide the specific service requested by the user;
- [redacted]
- Lenses are designed with privacy, safety, and inclusion considerations in mind [redacted]

Functionality

Lens functionality, as a general matter, includes the following:

- Lenses are accessed through the Lens Carousel within the Snapchat camera, through the Search interface, and through Lens Explorer;
- a Snap created with a Lens can be shared, published or saved consistent with the rules applicable to the destination surface;
- a Lens can be reported by a user using the in-app reporting menu, including under the illegal content reporting option (further described in Section 5.4.6);
- Sponsored Lenses are subject to the advertising review framework described in Section 5.7.

Snapchat+ features relevant to Lenses

Lens-related subscription features include Lens+, which is available as a standalone subscription and as part of Snapchat+ Platinum[12]. Lens+ provides access to exclusive generative-AI-powered

[12] https://help.snapchat.com/hc/en-us/articles/38227175820564-What-is-the-Snapchat-Lens-Monthly-Plan.

[redacted]

Lenses, subject to applicable monthly generation allowances. Snapchat+ Platinum additionally provides an experience without Lens ads.

Generative AI features relevant to Lenses

Lenses incorporate generative-AI tooling within the Lens Studio creation pipeline and within certain Lens experiences delivered through the Snapchat camera. Snap operates a set of safety measures specific to generative-AI Lenses, including:

- input and output scanning for harmful content using a combination of internal and external models;
- [redacted]
- contextual identification of generative-AI functionality, as applicable, through a sparkle icon, disclaimer, Context Card or tooltip[13];
- application, where supported, of a Snap Ghost with sparkles watermark to certain generative-AI visual content when exported or saved;
- retention limits applied to input data used in generative-AI Lens flows (for example, input images for high-latency Lens generation are deleted within 48 hours);
- for generative-AI Lenses delivered through Lens Studio (including any third-party-tool integrations), application of the Generative Lens Tools Terms and proxy routing of any third-party requests.

The substantive risk implications of generative-AI content created or delivered through Lenses, and the corresponding mitigation measures, are addressed in the relevant Section 4 assessments and in Section 5.9 (Content Authenticity).

2.7 Advertising

What is Snap's advertising product?

Advertising is the principal commercial offering of Snap's services and the principal source of Snap's revenue. Advertising is delivered to users in scope content surfaces of Snapchat, including Spotlight, Discover, Snap Map and Lenses (in the case of Sponsored Lenses), as well as in private and limited-broadcast surfaces that are out of scope of this Report. Advertising is created and managed by advertisers through Snapchat Ads Manager and the Snap Ads API.

Advertising shown to users is, as a general matter, subject to the Snap Advertising Policies, the Community Guidelines, and the in-scope-service moderation framework described in Section 5.4 (Content Moderation). The advertising system, including the ad-review pipeline, advertiser

[13] https://values.snap.com/privacy/privacy-by-product.

[redacted]

controls and audience-selection mechanisms, is described in detail in Section 5.7 (Advertising Systems).

How does Advertising work?

Advertisements are submitted by advertisers through Snapchat Ads Manager or through the Snap Ads API. Submitted advertisements are reviewed through a combination of automated screening and human review. [redacted] Submission systems are configured to prevent the selection of sensitive data categories for targeting and to prevent the submission of personalized/activity-based targeted advertisements aimed at minors.

Advertisements that pass review are eligible for delivery. [redacted] The applicable thresholds and controls are described further in Section 5.7.

All paid advertisements on Snapchat are presented with an "Ad" marker. Sponsored Lenses are presented with a "Sponsored" marker. Organic commercial content posted by influencers, creators or other users is subject to the Commercial Content Policy, which requires appropriate disclosure; a "Paid Partnership" tag tool is made available for users to apply to such posts.

Ad formats

Advertising on Snapchat is delivered through the following principal formats[14]:

- **Single Image or Video Ads.** Full-screen video advertisements with sound. Single Image or Video Ads include single-image or single-video skippable ads, with optional attachments enabling actions such as visiting a website or installing an application.
- **Story Ads.** Branded tiles within the Discover section of the Stories tab, comprising either video or a series of three to twenty images.
- **Collection Ads.** Advertisements presenting four tappable tiles to showcase multiple products.
- **Dynamic Ads.** Advertisements drawing on a product catalogue, served through the ad-ranking system.
- **Commercials.** Non-skippable advertisements (for the first six seconds), of up to three minutes' duration, presented within Snapchat's curated content.

[14] https://forbusiness.snapchat.com/advertising/ad-formats.

[redacted]

- **Sponsored Snaps.** Sponsored messages (which are clearly labeled as ads) presented within or adjacent to the message feed.
- **Promoted Places.** Sponsored places of interest highlighted on Snap Map with a promoted pin.
- **Sponsored Lenses (AR Ads).** AR Lenses delivered as advertising. Subject to the Lens review framework described in Section 2.6 (Lenses), the advertising review pipeline described above, and the Sponsored Lens marker.

Tools for advertisers

Snap makes the following advertiser-facing tools available, alongside Ads Manager and the Ads API:

- **Custom List Audiences.** Permit advertisers to provide Snap with information about their existing customers to inform audience-building.
- **Snap Pixel.** A web-based measurement tool enabling advertisers to attribute conversion events to advertising delivery.
- **Conversion API.** A server-to-server measurement integration enabling advertisers to share conversion data with Snap for measurement and audience-building.
- **Advanced and Estimated Conversion.** Measurement enhancements supporting advertisers in attributing and modelling conversion outcomes.

Each of the above tools is subject to advertiser-facing terms and to Snap's privacy and data-handling controls. Further detail in Section 5.7.

Functionality

Advertising-surface functionality, as a general matter, includes the following:

- a Snapchatter can report an advertisement using the in-app reporting menu, including under the illegal content reporting option (further described in Section 5.4.6); reported advertisements are subject to human moderation review;
- [redacted]
- a Snapchatter in the European Union can disable personalised content recommendations and adjust personalised-advertising preferences via the "European Union Controls" settings (further described in Section 5.7);[15]
- advertisements that pass review and are subsequently identified as violating, including through user reporting, are removed from delivery and the advertiser is notified.

Material 2026 update: political advertising in the European Union

[15] https://values.snap.com/privacy/snap-ads/ads-control.

[redacted]

As of October 2025, and consistent with the EU's Political Advertising Regulation, Snap has With effect from 1 October 2025, Snap no longer permits political advertising in the European Union as defined under Regulation (EU) 2024/900 on the transparency and targeting of political advertising. Accordingly, political advertising was not an active advertising-delivery vector in the European Union during the portion of the assessment period following that date. Snap's Political Ads LibRAy continues to provide historical transparency regarding previously approved political advertisements.[16]

Snapchat+ features relevant to Advertising

Snapchat+ Platinum subscribers[17] receive a reduced-advertising experience that does not include Sponsored Snaps or Story or Lens ads. Other advertisements may continue to be displayed.

Generative AI features used by Advertising

Snap makes a number of generative-AI tools available to advertisers within the advertising creation pipeline (for example, AI-assisted creative generation within Ads Manager[18]). Generative-AI advertising content is subject to the same advertising-review framework as other advertising content, and where applicable, to the content-authenticity measures described in Section 5.9.

3. DSA Risk Assessment Methodology

In preparing this Report, Snap has refined the risk assessment methodology applied in its 2023, 2024 and 2025 Reports to continue the iterative development of its DSA systemic risk assessment framework and to further strengthen the evidentiary record underlying the conclusions reached under Articles 34 and 35 DSA. The refinements, described further in the sub-sections below, are:

I. [redacted]

[16] Snap Political and Advocacy Advertising Policies.

[17] https://help.snapchat.com/hc/en-us/articles/18520949571988-What-is-the-Snapchat-Platinum-Monthly-Plan.

[18] https://forbusiness.snapchat.com/blog/human-first-ai-enabled-snaps-latest-ads-innovations.

[redacted]

V. [redacted]

No change has been made to the foundational framework applied in previous Reports. Snap continues to apply a standard risk assessment methodology adapted from those used in other regulatory contexts, including the EU general risk assessment methodology for product safety, the ICO's Data Protection Impact Assessment (DPIA) guidelines, and Ofcom's guidance for completion of illegal harms and protection of children risk assessments under the UK Online Safety Act.

The methodology comprises the steps set out below.

3.1 Identification of Risks

As a first step, Snap identified potential systemic risks for each of the four categories outlined in Article 34(1) DSA:

- **Category 1** (Article 34(1)(a) / Recital 80): dissemination of illegal content or content that violates Snap's terms and conditions, including CSAM, illegal hate speech, terrorist content, adult sexual content, harassment and bullying, content glorifying self-harm, fraud and spam, and content relating to other illegal activities;
- **Category 2** (Article 34(1)(b) / Recital 81): negative effects on the exercise of fundamental rights, including human dignity, freedom of expression and information, private life, data protection, non-discrimination and freedom of religion, children's rights, consumer protection and property;
- **Category 3** (Article 34(1)(c) / Recital 82): negative effects on democratic and electoral processes, civic discourse and public security; and
- **Category 4** (Article 34(1)(d) / Recital 83): negative effects on public health, gender-based violence, minors and physical and mental well-being (including addictions).

2026 refinement - surface-level risk pathways. In addition to the harm-category risks above, this Report identifies a set of cross-cutting contact and solicitation risk pathways operating at the surface or feature level across in-scope services. These include, without limitation:

- [redacted]
- [redacted]
- [redacted]
- [redacted]

These pathways are assessed alongside the corresponding harm categories (in the ordinary course, Category 4 risks concerning Minors and Category 1 risks concerning CSAM and certain illegal activities). Surface-level articulation permits a clearer view of where existing mitigations are

[redacted]

specific to a surface rather than horizontal, and supports more targeted proportionality analysis under Article 35.

3.2 Likelihood Analysis

As a second step, Snap analysed the extent to which each identified risk is likely to occur on Snapchat's in-scope services. As a general matter, the prevalence of the identified risks on Snapchat is low, in part because of existing mitigations and the design of the relevant functionality. To support consistent prioritisation and monitoring, Snap applies three likelihood bands anchored to fixed, internally defined quantitative thresholds where Prevalence Testing or Violative View Rate data is available. These bands assess likelihood on Snapchat and do not involve comparison with other platforms or require a predetermined distribution of risks across the three bands.

Level	Description
Highest likelihood	[redacted]
Medium likelihood	[redacted]
Lowest likelihood	[redacted]

Because observed PVP and VVR data reflect the operation of existing mitigations, Snap treats those metrics as evidentiary inputs rather than as a complete measure of inherent likelihood. Where both PVP and VVR data are available and indicate different likelihood bands, or where neither metric is available, Snap determines the likelihood category by considering the totality of the available internal and external evidence and records any material methodological judgment in the relevant risk section. Snap also considers the reasonably foreseeable likelihood of the risk absent the effect of the relevant mitigations.

Inputs to the likelihood analysis include: (i) internal data, including Prevalence Testing (PVP), Violative View Rate (VVR), illegal or violating content reporting data and complaint data; (ii) input from Snap's Safety Advisory Board and Council for Digital Well-Being; and (iii) external information, including academic research, reports from civil-society organisations and NGOs, government and regulator publications, and press reporting. Where internal data is required, it is

[redacted]

obtained from the functions responsible for maintaining it-for example, Customer & Review Operations with respect to complaint data.

As part of the continued iterative development of its methodology, Snap has introduced three voluntary enhancements to the likelihood analysis. These enhancements supplement, rather than replace, the framework applied in prior Reports and are intended to strengthen the evidentiary basis for the assessment.

(a) External data as a supplementary input where internal data is structurally limited. [redacted] For these risks, Snap has drawn on publicly available survey data from Member State regulators and research bodies as a likelihood proxy, in addition to [redacted] and proactive detection outputs.

(b) Feature-level articulation for contact risks. Where a risk materialises through a specific functional pathway (e.g., Search, Find Friends, contact suggestion surfaces), likelihood is assessed at the pathway level in addition to any aggregate platform figure. This permits a more targeted proportionality analysis at the mitigation stage.

(c) Performance information relating to automated systems. Where an age-inference model, proactive detection mechanism or other automated system materially informs the likelihood analysis, Snap may include relevant performance information where that information is available and appropriate to interpreting the assessment. The nature and level of detail disclosed will depend on the relevance, reliability and interpretability of the information.

(d) Refinement of Prevalence Testing confirmation methodology. During preparation of the 2026 assessment, Snap refined the methodology used to calculate category-level PVP by distinguishing between an initial-review determination and a confirmed violation following the applicable double-review process. The double-review methodology excludes sampled content initially identified as potentially violating where the violation is not confirmed on subsequent review and therefore produces a more targeted measure of confirmed violations.

To preserve comparability with prior assessment periods, this Report uses the initial-review methodology for year-on-year PVP trend analysis, consistent with the methodology underlying the historical figures reproduced from prior Reports. The corresponding double-review confirmed-violation figures are presented as supplementary 2026 context in the relevant Section 4 assessments and in Section 6.4 and are not used to determine year-on-year direction. Snap intends to apply the refined double-review methodology to PVP reporting on a going-forward basis.

[redacted]

3.3 Severity Analysis

As a third step, Snap analyzed the severity of each identified risk by reference to the potential harm to individuals and to society. Because every identified risk carries at least significant potential harm industry-wide (which is why it has been identified), Snap applies a relative severity measure across three levels:

Harm classification	Description
Severe harm industry-wide	[redacted]
Serious harm industry-wide	[redacted]
Significant harm industry-wide	[redacted]

Severity is informed by consultation with internal experts (including Trust & Safety, Policy and Legal), Snap's Safety Advisory Board and Council for Digital Well-Being, relevant external research, and guidance from regulators, international bodies and NGOs. Where applicable, Snap takes account of the vulnerability of affected populations, in particular, minors, in its severity assessment, and may weight severity upward where a risk disproportionately affects a vulnerable cohort.

3.4 DSA Risk Factors

Consistent with Article 34(2) DSA, Snap considered whether and how each of the following service risk factors influences each identified risk on Snapchat's in-scope services:

- the design of recommender systems and any other relevant algorithmic system;
- content moderation systems;
- the applicable terms and conditions and their enforcement;
- systems for selecting and presenting advertisements; and
- data-related practices.

Snap has also analyzed the influence of the general risk factors identified in Article 34(2):

- intentional manipulation of the service, including inauthentic use or automated exploitation;

[redacted]

- amplification and potentially rapid and wide dissemination of illegal content and of information incompatible with Snap's terms and conditions; and
- specific regional or linguistic aspects, including where specific to a Member State.

[redacted].

3.5 Overall Potential Risk Prioritization

As a fifth step, and before taking account of the effect of existing mitigations, Snap determined an overall inherent risk prioritisation for each identified risk, taking into account the likelihood and severity assessments described above. This prioritisation informs the subsequent assessment of the reasonableness, proportionality and effectiveness of the mitigations in place under Article 35(1) DSA.

Snap uses the following matrix as a guide. The matrix is an approximation: Snap determines the overall prioritization on a case-by-case basis and, where circumstances warrant, may depart from the matrix output. Any such departure is identified and explained in the relevant individual risk section.

Harm classification industry-wide	Lowest likelihood	Medium likelihood	Highest likelihood
Severe harm industry-wide	**Level 3**	**Level 1**	**Level 1**
Serious harm industry-wide	**Level 3**	**Level 2**	**Level 1**
Significant harm industry-wide	**Level 3**	**Level 3**	**Level 2**

Level 1 risks receive the highest prioritisation for mitigation and monitoring resources; Level 2 and Level 3 risks are prioritised in the ordinary course, subject to ongoing review. These levels reflect inherent risk prior to consideration of the effect of mitigations.

[redacted]

3.6 Snap's Mitigations

As a sixth step, Snap considered the mitigations in place to address each identified risk. Mitigations are mapped to the non-exhaustive list in Article 35(1)(a) to (k) DSA, as set out in the cross-reference table in Section 5.

[redacted]

I. **Reasonableness -** whether the mitigation is reasonable having regard to the nature of the risk, the population affected (including minors), and Snap's operational and financial capacity;
II. **Proportionality -** whether the mitigation is proportionate having regard to the overall potential risk prioritization and the impact of the mitigation on fundamental rights, including freedom of expression, privacy and data protection; and
III. **Effectiveness -** whether the mitigation is effective having regard to internal and external evidence of performance.

[redacted]

3.7 Conclusions

As a seventh step, after assessing the mitigations in place, Snap evaluated the residual risk associated with each identified risk and drew conclusions on whether the combination of mitigations is reasonable, proportionate and effective for purposes of Article 35(1) DSA.

In formulating proposed improvements, Snap considers what would be reasonable and proportionate by reference to: (a) the inherent and residual risk assessments; (b) relevant operational and implementation considerations; (c) relevant industry standards; and (d) regulator guidance where available. Further information regarding Snap's governance, monitoring and remediation-tracking framework is set out in Section 6.

4. DSA Risk Assessment Results

This Section records the results of the risk assessment of Snapchat's in-scope services carried out under Article 34 DSA. The assessment was conducted in accordance with the scope described in Section 2 and the methodology described in Section 3. Each identified risk has been reassessed for this Report, drawing on internal data, Digital & Data Impact Assessments (DDIAs),

external research, and information arising from Snap's engagement with the European Commission and national authorities during the assessment period.

The risks assessed affect a range of individuals, including Snapchat's user community, victims of crime, members of the public, and the content moderators who review content on Snapchat. Where applicable, impacts that extend beyond direct users, including impacts on the well-being of moderators — are identified within the relevant risk section within this document.

Consistent with Article 34(1) DSA, the identified risks are organised into four categories:

- **Category 1** - dissemination of illegal content or content that violates Snap's terms and conditions (Article 34.1(a); Recital 80), addressed in Section 4.1;
- **Category 2** - negative effects on the exercise of fundamental rights (Article 34.1(b); Recital 81), addressed in Section 4.2;
- **Category 3** - negative effects on democratic and electoral processes, civic discourse and public security (Article 34.1(c); Recital 82), addressed in Section 4.3; and
- **Category 4** - negative effects on public health, gender-based violence, minors, and physical and mental well-being including addictions (Article 34.1(d); Recital 83), addressed in Section 4.4.

For each identified risk, this Section records (i) a harm description; (ii) the likelihood assessment; (iii) the relative severity assessment; (iv) analysis of the Article 34(2) service and general risk factors; (v) the overall potential risk prioritization under the matrix set out in Section 3.5; (vi) the mitigations relied upon, with cross-reference to the detailed mitigation descriptions in Section 5; and (vii) Snap's conclusion for that risk, taking account of the nature and extent of the residual risk remaining after application of the mitigations, on whether the mitigations in place are reasonable, proportionate and effective under Article 35(1) DSA and whether any further measures or ongoing monitoring are warranted.

4.1 Category 1 - Dissemination of Illegal Content or Content that Violates Snap's Terms and Conditions

This Section 4.1 sets out Snap's assessment of the risk of dissemination of illegal content, and of content that is incompatible with Snap's terms and conditions (including the Community Guidelines), on Snapchat's in-scope services. The scope of Category 1 includes, in particular, the illegal content identified in Recital 80: child sexual abuse material (CSAM); illegal hate speech; and criminal offences and the conduct of illegal activities, including the sale of prohibited products or services, dangerous or counterfeit products, and illegally-traded animals.

In assessing each Category 1 risk, Snap considered the general risk factors set out in Article 34(2) DSA, in particular (i) the extent to which the risk is influenced by intentional manipulation of the

[redacted]

service, including by inauthentic use or automated exploitation; (ii) the extent to which Snapchat's in-scope services allow for amplification and potentially rapid and wide dissemination of the relevant content; and (iii) specific regional or linguistic aspects, where applicable.

Category 1

Consistent with the methodology set out in Section 3, the Category 1 assessment in this Report includes the following refinements:

- [redacted]
- **Mitigation effectiveness metrics.** [redacted], the Article 35(1) effectiveness analysis in the corresponding risk section is supported with quantitative performance metrics, subject to applicable confidentiality and security considerations.
- **Non-coercive criminal recruitment.** Non-coercive criminal recruitment of minors is identified as a distinct risk pathway and assessed alongside - rather than consolidated into - Snap's existing human-trafficking analysis. Because this risk has been characterised as primarily a minor-safety concern, the detailed assessment is located in Category 4 (Section 4.4.3, Negative Effects on Minors), with a cross-reference from Section 4.1.12 (Other Illegal Activities).

Summary of Category 1 determinations

The table below summarises Snap's assessment of likelihood, harm classification, overall potential risk prioritization, and the conclusion on mitigations under Article 35(1) DSA for each Category 1 risk. As a general matter, where a specific risk poses both (i) a risk of significant damage to the physical or emotional well-being of Snapchatters and (ii) an imminent, credible risk of severe harm - including threats to human life, safety and well-being - Snap treats that risk as severe harm and assigns a Level 1 overall risk prioritization.

Risk	Likelihood	Harm classification	Risk prioritization	Conclusion (Art. 35(1) DSA)
4.1.1 Dissemination of CSAM	Lowest	Severe	**Level 1***	Reasonable, proportionate and effective
4.1.2 Dissemination of Illegal Hate Speech	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective

4.1.3 Sale of Prohibited Products - Drugs	Lowest	Severe	**Level 1***	Reasonable, proportionate and effective
4.1.3 Sale of Prohibited Products - Weapons	Lowest	Serious	**Level 2***	Reasonable, proportionate and effective
4.1.3 Sale of Prohibited Products - Other	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective
4.1.4 Dissemination of Terrorist Content	Lowest	Serious	**Level 2***	Reasonable, proportionate and effective; under close monitoring
4.1.5 Dissemination of IP-Infringing Content	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective
4.1.6 Adult Sexual Content - Sexual Crimes	Lowest	Serious	**Level 2***	Reasonable, proportionate and effective
4.1.6 Adult Sexual Content - Other	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective; under monitoring
4.1.7 Harassment and Bullying	Lowest	Serious	**Level 2***	Reasonable, proportionate and effective
4.1.8 Self-Harm, Suicide and Eating Disorders	Lowest	Serious	**Level 2***	Reasonable, proportionate and effective
4.1.9 Violent or Dangerous Behaviour	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective
4.1.10 Harmful False Information	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective
4.1.11 Fraud and Spam	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective

4.1.12 Other Illegal Activities	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective

**Case-by-case upward adjustment from the default Section 3.5 matrix output, as explained in the corresponding risk section within this document.*

Note: The determinations in the table above reflect the outputs of the individual risk assessments in Sections 4.1.1 through 4.1.12 below. Where a determination has changed relative to the 2025 Report, that change is identified and explained in the corresponding risk section.

Each Category 1 risk is addressed in turn in the sub-sections that follow.

4.1.1 Dissemination of Child Sexual Abuse Material

This Section 4.1.1 sets out Snap's assessment of the risk of dissemination of child sexual abuse material (CSAM), and of related child sexual exploitation and abuse (CSEA) conduct, on Snapchat's in-scope services.

Harm Description

CSEA refers to content and activity involving the exploitation of a vulnerability, trust relationship, or imbalance of power to coerce, manipulate or deceive a minor (a person under the age of 18) into sexual activity or activity for the sexual gratification of another, online or in person, including grooming and the dissemination of child sexual abuse material (CSAM).

- **Grooming** refers to building trust, establishing a relationship and developing an emotional connection with a minor with a view to manipulating, exploiting or abusing them. It includes communication or behaviour that attempts to persuade or coerce a minor with the intent of sexual abuse or exploitation, or that leverages fear or shame to keep a minor silent.
- **CSAM** includes still images, videos, and illustrated, computer-generated or other realistic depictions, and live streaming of a child in a sexually explicit context or engaging in sexually explicit acts. CSEA may involve the generation, possession, distribution, promotion, or solicitation of such material, including both "known CSAM" (previously hashed and catalogued by organisations such as NCMEC) and "new CSAM" (not previously hashed or catalogued), whether authentic or AI-generated. It may also involve the sharing or promotion of hyperlinks that direct users to CSAM hosted off-platform, including via CSAM-hosting sites, darknet marketplaces, cloud drives, encrypted messaging channels, file dumps, and evasion techniques such as obfuscated links or URL shorteners.

CSEA also covers: accounts dedicated to sexually harassing, intimidating or extorting minors; teens or adults requesting, encouraging, enticing or coercing minors to engage in sexual acts or produce sexual images; content facilitating sex acts with a minor; content promoting the sale of

[redacted]

commercial sex with a minor; and threats to share, exploit or expose a minor's intimate images or videos without consent (sextortion), whether by an adult or by another minor.

Content involving minors that does not reach the CSEA threshold but is nevertheless inappropriate is also prohibited under Snap's terms and Community Guidelines.

Likelihood

Likelihood has been considered in accordance with the methodology set out in Section 3. On the basis of the evidence below, CSEA is assessed at the **lowest likelihood** category for Snapchat's in-scope services.

Internal data

Prevalence Testing indicates that Child Sexual Exploitation remained a low-prevalence category on Snapchat's in-scope services[19] during the 2026 assessment period. Using the methodology retained for year-on-year comparability in Section 6.4, Child Sexual Exploitation PVP was [redacted] for the 90-day period ending 30 June 2026, compared with [redacted] recorded as at 30 April 2025. This represents an increase under the comparable initial-review methodology, while remaining low in absolute terms.

The increase should be considered alongside the broader development of Snap's CSEA detection framework during the assessment period. Snap expanded its ability to identify and block CSEA-related material, including through additional partner-sourced URL signals and other proactive detection measures. These improvements provide greater visibility into CSEA-related activity and strengthen Snap's ability to identify and action relevant material. They do not, however, provide a basis for attributing the observed PVP movement to any single cause.

Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP for the 90-day period ending 30 June 2026 was [redacted]. The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction. Child Sexual Exploitation remained among the lowest-prevalence categories, and Snap continues to monitor this category through Prevalence Testing, enforcement data and the CSEA-specific detection and moderation controls described below.

Where CSEA content was identified proactively or following a report, the median turnaround time from detection to final action was 3 minutes during H2 2025, as reported in Snap's H2 2025 global Transparency Report.

Six-month average VVR data for EU users and EU minors during the period from 1 January to 30 June 2026 are set out below. The table provides a harm-specific surface breakdown for Spotlight, Public Story and Maps.

[19] Spotlight, Discover (Public Stories in VVR), Snap Map.

[redacted]

Enforcement reason	Spotlight VVR	Public Story VVR	Maps VVR
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

The above table is a harm-specific extract from the broader VVR dataset recorded in Section 6.4. Section 6.4 contains the complete available view-category breakdown and explains the methodology and scope of the VVR measure. The presentation above is limited to Spotlight, Public Story and Snap Map content, the principal public-content surfaces for which the relevant CSEA VVR is separately presented in this section.

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. The Lens review taxonomy does not include a CSEA- or CSAM-specific rejection category. During H1 2026, [redacted] Lens submissions were rejected under the broader category "Includes depictions of nudity or sex acts;" this figure is not limited to content involving minors and should not be interpreted as the number of CSEA- or CSAM-related Lens rejections.

External evidence

External research and reporting during the assessment period is consistent with the internal-data conclusions above:

- **Finnish NGO Protect Children study**. The February 2024 study by the Finnish NGO Protect Children, funded by the Tech Coalition and Safe Online, ranked Snapchat lowest among the social media platforms identified in the study as being used to search for, view or share CSAM (10%), compared to Instagram (29%), Twitter/X (26%), Discord (23%), TikTok (21%), Facebook (20%), YouTube (18%) and Reddit (17%)[20].
- **Internet Watch Foundation (IWF) Annual Report**. The Internet Watch Foundation (IWF) Annual Report 2024 records a continued year-on-year increase in URLs and direct reports containing CSAM at the industry level, and a sustained concentration of CSAM-hosting services in Europe (62% of global cases in 2024)[21].
- **NCMEC's 2024 CyberTipline**. NCMEC's 2024 CyberTipline data records 20.5 million reports of suspected child sexual exploitation, compared with 36.2 million in 2023,

[20] https://www.suojellaanlapsia.fi/en/post/tech-platforms-used-by-online-child-sexual-abuse-offenders-1.

[21] https://www.iwf.org.uk/annual-data-insights-report-2024/data-and-insights/reports-assessment/.

[redacted]

alongside increases in online enticement, financial sextortion and reports involving generative AI[22].

Generative AI

Generative-AI-created CSAM is an acknowledged industry-level concern, reflected in the NCMEC data referenced above. The assessment has not identified generative-AI-created CSAM as a material issue on Snapchat's in-scope services during the assessment period. This area remains under monitoring, including through the Generative AI DDIA.

Likelihood conclusion

Taking into account the PVP and VVR data, the low absolute prevalence of CSEA, external benchmarking, and the feature-level analysis above, the likelihood of CSEA on Snapchat's in-scope services remains in the **lowest likelihood category**. Under the methodology retained for year-on-year comparability, Child Sexual Exploitation PVP increased from [redacted] as at 30 April 2025 to [redacted] for the 90-day period ending 30 June 2026. Snap treats this as an increase in the comparable PVP metric, while noting that it occurred during a period in which Snap also expanded its ability to identify and action CSEA-related material. This included the introduction of partner-sourced CSE URL blocking through the Internet Watch Foundation in April 2026 and Lantern URL ingestion in June 2026, alongside Snap's existing hash- and content-based detection systems. The available data does not allow Snap to attribute the PVP movement to a single cause, including whether it reflects a change in underlying prevalence, improved identification and classification of CSEA-related material, or a combination of factors.

Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP was [redacted] Considered together with the very low VVR results and the other evidence set out above, these results do not alter the existing lowest-likelihood determination. The category will continue to be monitored through the controls and ongoing risk-detection processes described below.

Severity

CSEA is assessed as a risk of severe harm industry-wide. This conclusion is unchanged from prior Reports and reflects (i) the nature of the conduct itself as among the most serious crimes against children, and (ii) the documented lifelong consequences for victims, including where the abuse is facilitated or distributed online.

Relevant external sources bearing on severity include:

[22] https://www.missingkids.org/blog/2025/ncmec-releases-new-data-2024-in-numbers.

[redacted]

- The Canadian Center for Child Protection's 2017 International Survivors' Survey found that almost 70% of respondents worried about being recognised because of recordings of their abuse, while 85% anticipated needing ongoing or future therapy[23].
- InHope guidance on grooming, identifying anxiety, depression, post-traumatic stress and suicidal ideation as likely long-term mental-health consequences for victims, including where no personal contact has occurred[24].
- IWF's 2024 data indicates that 91% of assessed URLs included at least one image or video of a "self-generated" nature, with children aged 11–13 and 7–10 the most commonly represented age groups. Earlier IWF data recorded an increase in instances involving children aged 7–10 from approximately 8,000 in 2020 to approximately 27,000 in 2021[25].
- The European Commission's *EU Strategy for a More Effective Fight Against Child Sexual Abuse* describes child sexual abuse as a particularly serious crime with wide-ranging and serious lifelong consequences, while the Council of Europe similarly recognises the far-reaching, long-term consequences of online CSEA[26].

Severity is treated as elevated for minors within the affected population, consistent with the methodology set out in Section 3.3.

DSA Risk Factors

In accordance with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services. While the factors could, in principle, influence this systemic risk, their overall impact is limited in view of the mitigations described in Snap's Mitigations below and in Section 5.

Design of recommender systems and other algorithmic systems	Recommender systems are designed not to recommend CSEA. There are no CSEA interest categories, and no CSEA content is surfaced through ranking or personalisation systems. Further detail is provided in Section 5 (Algorithmic Systems).
Content moderation systems	Specific proactive and reactive moderation procedures to prevent and remove CSEA are in place, including [redacted] The Article 35(1) effectiveness analysis is supported by quantitative

[23] https://protectchildren.ca/en/resources-research/survivors-survey-results/.
[24] https://inhope.org/articles/the-impact-of-online-grooming.
[25] Three-fold increase of abuse imagery of 7-10-year-olds as IWF detects more child sexual abuse material online than ever before.
[26] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52020DC0607.

[redacted]

	performance metrics (see the Effectiveness Metrics sub-section below). Further detail is provided in Section 5 (Content Moderation).
Applicable terms and conditions and enforcement	The Community Guidelines (which form part of the terms of service) prohibit CSEA for all users, and are enforced in accordance with the severity of the harm. Account-level enforcement, content removal and, where applicable, reporting to the National Center for Missing & Exploited Children (NCMEC) are applied. Further detail is provided in Section 5 (Terms).
Systems for selecting and presenting advertisements	Advertising policies prohibit content depicting or promoting CSEA, and advertisements are subject to pre-publication review processes before publication. The assessment has not identified advertising as a material vector for the dissemination of CSEA on in-scope services. Further detail is provided in Section 5 (Advertising Systems).
Data-related practices	Privacy-by-design practices include the PASS (Privacy Assessment System) review, and DDIAs covering Snaps & Chats, Public Profiles, Snap Map, Public Content, Lenses, Family Center and Generative AI. DDIA outputs feed into the likelihood and mitigation analysis in this section.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	In-scope services have design and moderation features that limit the amplification of CSEA content: [redacted]. These features are further described in Section 5 (Snapchat Design and Function; Content Moderation).
Regional and linguistic aspects	The service is available in multiple EU languages, Terms are published in all EU languages, and moderators operate across multiple EU languages. For EU operations, Abusive Language Detection uses English terminology as a baseline together with terminology from 14 additional EU languages, with terms and rules tailored by context and region.

[redacted]

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5 and the case-by-case adjustment permitted by the methodology, CSEA is assessed at **Level 1** overall potential risk prioritization. This reflects the severe harm industry-wide classification and the case-by-case adjustment to Level 1 applied to risks where the conduct carries an imminent, credible risk of severe harm notwithstanding a lowest likelihood occurrence profile. There is no change in this assessment from prior Reports.

[redacted]

Snap's Mitigations

The table below maps the mitigations in place for this risk to the non-exhaustive list in Article 35(1)(a)–(k) DSA. As mitigations are described in full in Section 5, the descriptions below are illustrative rather than exhaustive; detailed descriptions are cross-referenced by sub-section.

Snapchat Design and Function - Art. 35(1)(a)	Applies. Design features limit unwanted contact with minors, including: ([redacted] (see Section 5.1.3). Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines prohibit CSEA for all users. Account-level enforcement is applied where CSEA is identified on the platform, and reports are submitted to NCMEC as applicable, which coordinates with law enforcement in the United States and elsewhere. Further detail in Section 5.2.

[redacted]

Moderation - Art. 35(1)(c)	Applies. Proactive and reactive moderation procedures operate across in-scope services. Proactive scanning of Public Stories, Spotlight and Discover submissions uses [redacted] at upload; Public Stories, Spotlight and Discover are additionally evaluated via [redacted]. Proprietary signal-based tools address sextortion and other sexual harms against minors. Reactive reporting is available in-app and via the Support Site. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Recommender systems are configured not to surface CSEA: no CSEA interest categories exist, and all Spotlight submissions require automated approval before public visibility, [redacted] Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising policies prohibit CSEA, and advertisements are subject to pre-publication review. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing, VVR monitoring, enforcement data and Transparency Report outputs are used to monitor CSEA risk on a continuous basis. CSEA-specific queues and a specialised human-review workflow (including double-review for certain categories) are maintained. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Notices concerning CSEA submitted by designated trusted flaggers under Article 22 DSA receive the priority treatment described in Section 5.10. [redacted] Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Participation in industry coordination groups and signal-sharing programmes includes the [redacted]. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. Guidance on terms, harms, moderation and reporting is published in the Safety Center. In-app resources address sexual content (including CSEA), reporting pathways, and support for users affected by financial sextortion. Further detail in Section 5.3.

[redacted]

Protection of Minors - Art. 35(1)(j)	Applies. Measures include (i) restrictions on minors' contact with non-friends; (ii) in-chat safety prompts on messages [redacted] (iii) Family Center for parental oversight of recent conversation activity and friend additions; and (iv) educational content via the Snapchat Family Safety Hub and the parents site. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Partially applies. An AI sparkle icon is applied to certain generative-AI outputs produced in Snap tools (for example, Bitmoji Backgrounds). Generative-AI-created CSAM is monitored as an emerging risk, consistent with NCMEC reporting trends; the assessment has not identified a material incidence of generative-AI-created CSAM on in-scope services during the assessment period.

Effectiveness metrics

Consistent with the Article 35(1) proportionality frame, the effectiveness of the mitigation set is assessed using quantitative performance information where available. The principal metrics used or being developed for this purpose are summarised below.

- [redacted]
- **Partner-sourced CSE URL blocking.** Since 6 April 2026, Snap ingests and blocks known child sexual exploitation [redacted]
- **Novel CSAM detection** [redacted] Precision and recall at deployed operating point; monthly volume of auto-enforcement and monthly volume escalated to human review.
- **Sextortion and sexual-harm signal-based detection.** Volume of sextortion-tagged cases routed to Trust & Safety; median time to enforcement; proportion enforced.

[redacted]

- **Abusive Language Detection** [redacted] Dictionary coverage across EU languages; monthly volume of auto-rejections; sample-based false-negative rate.
- **Contact-restriction controls** [redacted]
- **Reactive reporting pipeline.** Volume of CSEA reports received in-app and via Support Site; median time to first action; onward NCMEC reporting volume.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of CSAM and related CSEA conduct on Snapchat's in-scope services are reasonable, proportionate and effective.

The residual-risk items identified in this risk section are recorded in the governance and action-tracking framework described in Section 6, with assigned owners and review cadences.

4.1.2 Dissemination of Illegal Hate Speech

This Section 4.1.2 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of illegal hate speech on Snapchat's in-scope services.

Harm Description

Snap describes hate-based content as content that demeans, defames or promotes discrimination towards an individual or a group of individuals on the basis of race, colour, caste, ethnicity, national origin, religion, sexual orientation, gender, gender identity, disability, veteran status, immigration status, socio-economic status, age, weight, or pregnancy status. This includes, without limitation, references to people that are dehumanising, that compare humans to animals on the basis of these traits or categories, the valorisation of perpetrators (or the denigration of the victims) of hateful atrocities (such as genocide, apartheid or slavery), and the promotion of hate symbols.

Examples of content that, in the ordinary course, falls within the scope of this category include:

- glorification of members of a hate group, including images that are not newsworthy or lawful political speech or social commentary;
- sharing of manifestos of hate groups or individuals acting under a hate ideology in a context that is supportive or that lacks contextualising or critical commentary;
- slurs used outside of self-reference, including emojis used to represent slurs or racial stereotypes;
- attacks on an individual or group on the basis of a protected characteristic;

[redacted]

- symbols of hate groups, or hate symbols intended to convey their hateful meaning;
- memes that target, shame or call for discrimination, segregation, violence or hate against a protected class;
- memes that deny historical acts of genocide or other atrocities, or which denigrate the victims of such tragedies;
- intentional deadnaming or misgendering of transgender or non-binary individuals.

The category covers both content that is illegal under EU or Member State law (including under Council Framework Decision 2008/913/JHA on combating racism and xenophobia, and corresponding national implementations) and content that is incompatible with the Community Guidelines but not necessarily illegal in every Member State. The Community Guidelines apply uniformly across in-scope services and are enforced consistent with that uniform standard.

Likelihood

Internal data

Prevalence Testing indicates that Hateful Content remained a low-prevalence category on Snapchat's in-scope services[27] during the 2026 assessment period. Using the methodology retained for year-on-year comparability in Section 6.4, Hateful Content PVP was [redacted] for the 90-day period ending 30 June 2026, compared with [redacted] recorded in the 2025 Report. This reflects a continued downward movement in the prevalence of policy-violating hateful content under the comparable initial-review methodology. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP for the 90-day period ending 30 June 2026 was [redacted]. The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction.

The PVP category covers content prohibited under Snap's Hateful Content policy and is used as an indicator relevant to the assessment of illegal hate speech. It is not limited only to content determined to be illegal under the law of a particular Member State.

Where Hate Speech content was identified proactively or following a report, the median turnaround time from detection to final action was 1 minute during H2 2025, as reported in Snap's H2 2025 global Transparency Report.

Six-month average Violative View Rate data for EU users and EU minors across the in-scope public surfaces of Spotlight, Public Stories and Maps during the first half of 2026 are set out below:

Cohort	Spotlight VVR	Public Story VVR	Maps VVR

[27] Spotlight, Public Story, Maps, Saved Story, Communities.

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

These figures remain within the lowest likelihood category under the methodology in Section 3.

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. For Lenses, during H1 2026, submissions were rejected for reasons relevant to hate-based or protected-group content, including Harassment/Hate Speech [redacted] inappropriate or offensive content targeting a particular group [redacted] disability-based content [redacted] and skin-tone-based content [redacted]

External evidence

External evidence specific to Snapchat under the EU Code of Conduct on Countering Illegal Hate Speech Online+ supports the low-likelihood determination. In the first monitoring exercise under the revised Code+, conducted from early November to mid-December 2025, all twelve signatories agreed to participate. However, only Facebook, Instagram, TikTok, X and YouTube received relevant notifications from the independent monitoring reporters; Snapchat was not among the platforms that received a relevant notification.[28]

This outcome is consistent with the earlier monitoring record under the original Code. In the seventh monitoring exercise, conducted in 2022, 36 specialist organisations across 21 Member States submitted 3,634 notifications to participating companies. Snap received no monitoring notification for the fifth consecutive year.[29]

- **Regulatory relevance of the evidence.** The Commission and European Board for Digital Services incorporated the revised Code+ into the Article 45 DSA framework on 20 January 2025. For VLOP signatories, including Snapchat, adherence is subject to independent annual auditing, and the Commission identifies the Code as a relevant benchmark for assessing systemic-risk mitigation under Article 35 DSA.[30]
- **Appropriate evidentiary limitation.** The monitoring outcome does not demonstrate the complete absence of illegal hate speech on Snapchat. It provides a targeted external indicator that specialist monitoring reporters did not identify and notify Snapchat of

[28] European Commission, 'First Results Published under the Revised Code of Conduct on Countering Illegal Hate Speech Online +' (10 April 2026).
[29] European Commission, Factsheet – 7th Monitoring Round of the Code of Conduct (1 December 2022); European Commission, 'First Results Published under the Revised Code of Conduct on Countering Illegal Hate Speech Online +' (10 April 2026).
[30] European Commission, 'The Code of Conduct on Countering Illegal Hate Speech Online +' (20 January 2025).

qualifying content during the defined monitoring window, considered alongside Snap's PVP, VVR and enforcement evidence.

Likelihood conclusion

On the basis of the internal Prevalence Testing data, the consistent downward PVP trend, and the external Code-monitoring outcomes, the likelihood of dissemination of illegal hate speech on Snapchat's in-scope services is assessed at the **lowest likelihood** category. There is no change in this determination from the 2025 Report.

Severity

Hate speech is a category of serious societal concern in the European Union, recognised by the EU Charter (Articles 1 and 21), the European Convention on Human Rights (Article 14), and Council Framework Decision 2008/913/JHA, among other instruments. The Council of Europe has noted that hate speech in online contexts can produce serious consequences for human dignity and contribute to discrimination and violence on the basis of protected characteristics.

The severity of negative effects associated with this category is assessed as **significant harm industry-wide**, consistent with prior Reports. As a general matter, and in line with the framework recorded in Section 3.5, any specific instance of hate-based content that presents an imminent, credible risk of severe harm - including incitement to physical violence on protected grounds - is treated as a Level 1 overall risk and handled accordingly.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services. The relevance of each factor to this systemic risk is considered below. The corresponding mitigations are assessed separately under Article 35.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend content that violates the Hate Speech provisions of the Community Guidelines; no interest categories corresponding to hate-based content exist. Hate-based content does not, in the ordinary course, achieve algorithmic distribution on Spotlight or Discover. Further detail in Section 5.6 (Algorithmic Systems).

[redacted]

Content moderation systems	Proactive and reactive moderation procedures operate to identify and remove hate speech across in-scope services. Automated detection [redacted] operates as part of the Search and contact-surface moderation framework described in Section 5.4.4. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines prohibit Hateful Content, Terrorism and Violent Extremism, including Hate Speech and content that demeans, defames or promotes discrimination or violence on protected grounds (including race, colour, caste, ethnicity, national origin, religion, sexual orientation, gender, gender identity, disability, veteran status, immigration status, socio-economic status, age, weight or pregnancy status). The use of hate symbols and the glorification of perpetrators of hateful atrocities are also prohibited. Further detail in Section 5.2 (Terms) and Section 5.5 (Enforcement).
Systems for selecting and presenting advertisements	The Advertising Policies prohibit hate-based content in advertisements. Advertisements are subject to pre-publication review using automated screening and risk-based human review. Reactive reporting pathways are available to users. Special-targeting models for housing, credit and employment (HCE) advertising exclude gender and age targeting to mitigate discrimination risk. Further detail in Section 5.7 (Advertising Systems).
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework described in Section 6.3.
Intentional manipulation (Article 34(2))	[redacted]

[redacted]

Amplification and rapid, wide dissemination	In-scope services are configured in a way that limits the amplification of hate-based content: short default retention for public stories (unless saved to Public Profile), the absence of an open recommender feed for unvetted user-generated content, pre-distribution moderation for content recommended to wide audiences on Spotlight and Discover, and exclusion of hate-based content from algorithmic recommendation. Limited Broadcast Content (Friends Stories, Community Stories, Private Stories) is, in the ordinary course, out of scope of this Report; however, proactive scanning is applied for severe-harm categories on these surfaces, as described in Section 5.4.3.
Regional and linguistic aspects	The service is available in multiple EU languages and the Terms are published in all official EU languages. ALD banned-word lists are maintained across multiple languages and categorised by context and region; moderation guidance is calibrated to Member State legal contexts where required (for example, in respect of categories of content that are illegal in specific Member States but not others, including specific symbols, denial of historical atrocities and specific protected-class slurs).

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, dissemination of illegal hate speech is assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of significant harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

[redacted]

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to include proactive moderation for hate speech across the public-content surfaces. The absence of an open recommender feed for unvetted user-generated content, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation for content reaching wide audiences combine to limit the amplification of hate-based content. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines prohibit Hate Speech and the use of hate symbols, and apply equally to all users including public figures, creators and Media Partners. Enforcement applies without regard to whether content is user-generated or AI-generated. Further detail in Section 5.2.
Moderation - Art. 35(1)(c)	Applies. Proactive moderation includes ALD scanning [redacted] multi-modal classifiers operating across media and text; [redacted] Reactive moderation is available through in-app reporting, with reports routed to human moderators trained on hate-speech policy. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend hate-based content; no corresponding interest categories exist. The Content Guidelines for Recommendation Eligibility set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising policies prohibit hate-based content. Special-targeting controls for HCE advertising mitigate discrimination risk in advertising delivery. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Hate Speech is conducted on a regular cadence, with results disclosed in this Report and in Snap's Transparency Reports. Further detail in Section 6.

[redacted]

Trusted Flaggers - Art. 35(1)(g)	Applies. [redacted]
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap is a signatory of the EU Code of Conduct on Countering Illegal Hate Speech Online (originally signed in 2018). The EU Code of Conduct was revised on 20 January 2025 to bring it into alignment with the DSA, and Snap has complied with the revised Code as part of its DSA compliance program. Snap also participates in the [redacted] and the Code of Conduct of the Freiwillige Selbstkontrolle Multimedia-Diensteanbieter (FSM) in Germany. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Hate Speech category explainer (in the Community Guidelines Explainer Series), the Safety Center, and Snap's Transparency Reports provide information on hate-speech policy, enforcement and prevalence. In-app reporting is available, with statements of reasons issued under Article 17 DSA in connection with enforcement decisions. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Mitigations for minors include the contact restrictions and Family Center oversight described in Section 5.8, and the minor-account default settings and onboarding refinements described in Section 5.1.3. [redacted] usernames are not displayed on Spotlight content created by minors. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines apply without regard to whether hate-based content is user-generated or AI-generated. Content-authenticity measures (watermarking, contextual identification) for generative-AI content are described in Section 5.9 and operate alongside the substantive prohibitions in the Community Guidelines.

Conclusion

[redacted]

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of illegal hate speech on Snapchat's in-scope services are reasonable, proportionate and effective.

Hate Speech continues to be one of the lowest-prevalence categories on Snapchat's in-scope services. The combination of the platform architecture [redacted] the proactive moderation suite [redacted] the Article 22 DSA trusted-flagger cooperation, [redacted] supports the conclusion that the mitigation framework operates effectively against this risk. There is no change in the overall conclusion from the 2025 Report.

4.1.3 Dissemination of Information Related to the Sale of Prohibited Products or Services

This Section 4.1.3 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of content related to the sale of prohibited or regulated products or services on Snapchat's in-scope services. The risk is sub-divided for severity-classification purposes into three categories: drugs (Severe industry-wide), weapons (Serious industry-wide), and other prohibited or regulated goods and services (Significant industry-wide), each assessed at lowest likelihood. Consistent with prior Reports, the overall risk prioritisation is recorded for each sub-category separately.

Harm Description

This harm covers content or activity that promotes, facilitates or enables the illegal supply, trade or solicitation of drugs, weapons, and other prohibited or controlled goods and services. It encompasses user-generated content, profile metadata, communications and interactions that:

- advertise or promote the sale of illicit drugs or controlled substances;
- offer firearms, knives, weapons or promote illegal weapons trafficking;
- enable or promote the trafficking of individuals, including content relating to the recruitment, transportation or exploitation of people through force, fraud or coercion;
- facilitate the sale or procurement of other goods and services prohibited by law (including counterfeit goods or documents, illegally-traded animals and similar prohibited categories);
- constitute the unauthorised promotion of regulated goods or industries, including gambling, tobacco and vape products, alcohol, and other regulated categories.

Snap prohibits any content that encourages or enables engagement in illegal transactions, or that constitutes the unauthorised promotion of regulated goods or industries. Further detail is set out

[redacted]

in the Snap Transparency Report Glossary[31] and in the Community Guidelines explainer on Illegal or Regulated Activities.

[redacted]

2026 Risk Assessment review

Snap's 2026 risk assessment process identified a number of areas in which existing mitigations warranted strengthening or extension, and in which methodological refinements could improve

[31] https://values.snap.com/privacy/transparency/glossary

[redacted]

the granularity and transparency of the overall risk analysis. The process drew on a range of external and internal sources, including a 2025 study by the Danish NGO Digitalt Ansvar[32] concerning Snapchat's handling of drug-related profiles and their recommendation to minor accounts. Snap reviewed the identified areas and implemented or strengthened mitigation measures where appropriate. The identified areas, Snap's responses, and cross-references to the relevant Section 5 sub-sections are set out below:

Text-moderation system coverage; [redacted] (Article 28(1))	Addressed in Section 5.4.4 (Search and contact surface moderation, 2026): [redacted] Outputs disclosed in Section 5.4.7 effectiveness metrics.
Broadcast moderation; reported profiles pipeline (Article 28(1))	Addressed in Section 5.4.5 (Product-Specific Moderation): audit of the reported-but-not-actioned profile pipeline; review of recommendation surfacing of flagged accounts to minors. Section 5.6 (Algorithmic Systems) addresses the recommendation-surface dimension specifically: [redacted]
Safety intervention surfaces; Heads Up expansion	Heads Up is an informational safety-support surface displayed at the top of Search results when users search for predefined drug-related terms. During the assessment period, localised Heads Up portals were available for the United Kingdom, France and the Netherlands, with the Netherlands portal launched in June 2026; users elsewhere in the EU were directed to the English-language US portal. Snap continues to explore additional localised versions.
Feature-level articulation of risk vectors; Search and Find Friends (Articles 34(1)–(2))	Addressed in this Section 4.1.3 (by recording Search and Find Friends as feature-level vectors for this risk in the Likelihood discussion below) and in Section 5.4.4 (which records the corresponding mitigations at the feature level).

[32] https://www.digitaltansvar.dk/aktuelt/snapchat-ignorerer-anmeldelser-af-danske-narkoprofiler.

[redacted]

Text-moderation system; operational scope and proportionality (Article 35(1))	Addressed in Section 5.4.3 (Detection systems table) and Section 5.4.4 (Search and contact surface moderation), which document the operational scope of ALD, multi-modal classifiers and signal-based detection; and in Section 5.4.7 (Effectiveness metrics), which records the disclosure of category-level precision, recall and false-positive metrics. The Article 35(1) proportionality analysis is recorded against the Section 6.1 Framework.
Proactive Detection; public-facing Stories (Article 35(1))	Addressed in Section 5.4.3 (Proactive moderation, including the Proactive Safety Detection System for severe-harm categories), Section 5.4.5 (Product-Specific Moderation, including the virality circuit-breaker on Spotlight), and Section 5.4.7 (Effectiveness metrics). The Digitalt Ansvar findings are addressed by the audit of the reported-but-not-actioned profile pipeline.
Safety intervention; Heads Up proportionality (Article 35(1))	Addressed in Section 5.1.3, Section 5.4.4 and this section. Heads Up is an informational safety-support surface displayed at the top of Search results when users search for predefined drug-related terms. During the assessment period, localised Heads Up portals were available for the United Kingdom, France and the Netherlands, with the Netherlands portal launched in June 2026; users elsewhere in the EU were directed to the English-language US portal. As a targeted Search intervention presented in response to relevant queries, Heads Up operates as a supplementary informational measure and supports the Article 35(1) proportionality analysis. Further detail is recorded in Section 6.1.5.

Likelihood

Internal data

Prevalence Testing indicates that content relating to prohibited or regulated products remained at a low level on Snapchat's in-scope services during the 2026 assessment period. Using the methodology retained for year-on-year comparability in Section 6.4, for the 90-day period ending 30 June 2026:

- Regulated Goods & Activities PVP was [redacted]
- Drugs PVP was [redacted]
- Weapons PVP was [redacted]; and
- Illegal Goods & Activities PVP was [redacted]

[redacted]

Compared with the results recorded in the 2025 Report, Regulated Goods & Activities decreased from [redacted] and Illegal Goods & Activities decreased from [redacted] while Drugs increased from [redacted] and Weapons increased from [redacted] The 2025 PVP metric for Weapons was limited to 'Weapon Sales' whereas the 2026 definition is broader and also includes content containing the display of weapons and brandishing. Comparing the same severe Weapon Sales measure year over year, it is flat. Notwithstanding those movements, each category remained within the lowest-likelihood range under the methodology in Section 3.

Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP figures for the 90-day period ending 30 June 2026 were [redacted] for Regulated Goods & Activities, [redacted] for Drugs, [redacted] for Weapons, and [redacted] for Illegal Goods & Activities at the displayed level of precision. These refined figures are presented as supplementary 2026 context and are not used to determine the year-on-year direction.

These categories provide the available PVP indicators relevant to the three sub-categories assessed for this risk. For purposes of interpreting the PVP results:

- **Regulated Goods & Activities** covers unauthorised content involving substances, goods or activities regulated by government authorities. This includes the sale or unauthorised promotion of regulated goods, such as tobacco and nicotine products, e-cigarettes and vapes, alcohol and financial products, and the promotion or solicitation of regulated activities such as gambling, prize lotteries and multi-level marketing practices.
- **Drugs** covers policy-violating content involving drugs and controlled substances, including attempts to buy, sell or distribute conditionally prohibited substances—such as cannabis-derived products, psilocybin, nitrous oxide and certain prescription or over-the-counter medicines—and strictly prohibited substances, such as heroin, cocaine, methamphetamine, fentanyl and commonly abused prescription medicines.
- **Weapons** covers policy-violating content promoting or facilitating the attempted purchase, sale or distribution of weapons, ammunition or explosives, including both lethal and non-lethal weapons and related goods.
- **Illegal Goods & Activities** is a residual PVP classification for policy-violating illegal goods or activities that are not classified within the Drugs, Weapons or Regulated Goods & Activities categories. Relevant policy examples include the attempted sale of counterfeit, pirated or stolen goods or documents. Other forms of illegal conduct, including organised crime and human trafficking, are separately defined within Snap's Trust and Safety policies.

[redacted]

These categories are separately classified within the Prevalence Testing taxonomy and should not be aggregated. The PVP categories are operational classifications used for prevalence measurement and do not independently determine whether particular content is illegal under the law of a specific Member State.

[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Six-month average Violative View Rate data for EU users and EU minors across the in-scope public surfaces of Spotlight, Public Stories and Maps during the first half of 2026 are set out below.

[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. For Lenses, during H1 2026, submissions were rejected for reasons including substance abuse involving drugs [redacted], tobacco [redacted] and alcohol [redacted], promotion of dangerous, harmful or illegal activity [redacted], and content inciting or glorifying violence or the use of weapons [redacted]

External evidence

Public-health authorities, law enforcement and regulators in the European Union have increasingly identified the migration of drug markets to digital platforms as an area of concern. EUDA[33] has reported that social media and instant-messaging applications are becoming increasingly prominent channels for the retail sale and supply of drugs. A 2023 study published in EUDA's document library examined social-media drug dealing in Denmark, Norway, Sweden and Germany, and reported ready access to drug dealers through Snapchat in Denmark, Norway and Sweden.

In addition, the 2025 study by the Danish NGO Digitalt Ansvar identified Snapchat-specific concerns relating to drug profiles remaining accessible after report and being recommended to minor accounts.[34]

The Digitalt Ansvar study also identified evasion and discoverability concerns involving explicit drug-related terminology in usernames and profiles findable through Search. Snap considered these findings as part of its review of text- and profile-level moderation. The corresponding mitigation measures are described in Section 5.4.4.

[redacted]

[33] Drivers and facilitators — new report unveils dynamics of illicit drug market.
[34] https://www.digitaltansvar.dk/aktuelt/snapchat-ignorerer-anmeldelser-af-danske-narkoprofiler.

[redacted]

[redacted]

Likelihood conclusion

On the basis of the low PVP and VVR results, the internal enforcement data, and the operational mitigations recorded above, the likelihood of dissemination of content relating to the sale of prohibited or regulated products on Snapchat's in-scope services is assessed at the **lowest likelihood category** for each of the three sub-categories (drugs, weapons, and other regulated/prohibited goods including vapes). There is no change in this determination from the 2025 Report. The mitigation framework is further strengthened by the changes recorded in the sub-section above.

Severity

The severity of harm associated with this risk varies materially across sub-categories:

Drugs

The dissemination of content facilitating the sale of illicit drugs or controlled substances is associated with severe harm industry-wide, including in particular the well-documented public-health emergency arising from the trafficking of synthetic opioids (notably fentanyl and counterfeit pills containing fentanyl). The severity classification reflects the imminent and credible risk to human life that flows from this category of content.

Weapons

The dissemination of content facilitating the sale of firearms, knives or other weapons is associated with serious harm industry-wide, recognising the risk to physical safety while distinguishing the category from drugs (where the risk pathway is more frequently a direct threat to life).

Other regulated and prohibited goods (including vapes)

The dissemination of content facilitating the sale or unauthorised promotion of other prohibited or regulated goods (including counterfeit goods, illegally-traded animals, gambling, alcohol, tobacco and vape products) is associated with significant harm industry-wide. [redacted]

[redacted]

As a general matter, and consistent with Section 3.5, any specific instance within any sub-category that presents an imminent, credible risk of severe harm — including a threat to human life, safety or well-being — is treated as a Level 1 overall risk and handled accordingly.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services. [redacted]

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend content concerning the sale of prohibited or regulated products or services; no corresponding interest category exists. [redacted]. Further detail in Section 5.6.
Content moderation systems	[redacted] The expansion of profile-level moderation in Section 5.4.4 strengthens detection [redacted] Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines prohibit (i) any use of Snapchat for content that is illegal in the user's jurisdiction or for any illegal activity, including buying, selling or facilitating sales of illegal or regulated drugs, weapons, contraband, counterfeit goods or documents, and (ii) the unauthorised promotion of regulated goods or industries, including gambling, tobacco and vape products, and alcohol. [redacted] Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	The Advertising Policies prohibit the advertising of vapes, vape products and other prohibited or regulated products that are not subject to a permitted promotion category. Pre-publication review applies to all advertising. Further detail in Section 5.7.

[redacted]

Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of prohibited-product content on in-scope services is constrained by: the absence of a marketplace function on Snapchat; the absence of an open recommender feed for unvetted user-generated content; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences on Spotlight and Discover; [redacted]
Regional and linguistic aspects	ALD banned-word lists are maintained across multiple languages and categorised by context and region. Cooperation relevant to Member State-specific risks includes Article 22 trusted-flagger notices and, separately, engagement with specialist organisations and law-enforcement bodies, including the National Crime Agency in the UK in respect of the sale of knives and drugs.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, the overall prioritisation for this risk is recorded by sub-category as follows:

[redacted]

[redacted]

[redacted]

The matrix-level positioning of each sub-category is shown below:

[redacted]

The determination for other regulated or prohibited goods corresponds to the matrix default for significant harm industry-wide combined with lowest likelihood and is therefore Level 3. For drugs and weapons, Snap applies the case-by-case deviation framework in Section 3.5. The default matrix output for both severe harm industry-wide combined with lowest likelihood and serious harm industry-wide combined with lowest likelihood is Level 3. Drugs are assessed at Level 1 overall potential risk prioritisation, reflecting the potential for imminent and credible threats to human life associated with illicit-drug distribution, including synthetic opioids and counterfeit pills containing fentanyl. Weapons are assessed at Level 2, reflecting the serious physical-safety risks associated with the unlawful sale and distribution of firearms, knives and other weapons. There is no change in these determinations from the 2025 Report.

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to limit the structural opportunity for prohibited-product distribution: no marketplace function; recommended content in public feed is moderated; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences. The minor-account default settings (Section 5.1.3) further limit minor exposure to prohibited-product content. The Heads Up intervention

[redacted]

	(Section 5.1.3) provides an informational safety-support surface at the top of Search results for predefined drug-related queries.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines prohibit illegal activities including the sale of illegal or regulated drugs, weapons, contraband and counterfeit goods, and prohibit the unauthorised promotion of regulated goods including tobacco and vape products. Enforcement applies severe-harm consequences and, in the ordinary course, retains content for an extended period to support law-enforcement investigation. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Proactive moderation includes ALD [redacted]; multi-modal classifiers [redacted] the Proactive Safety Detection System with auto-enforcement [redacted] and human review of content [redacted] The expansion of profile-level moderation in Section 5.4.4 addresses the Search and Find Friends vectors directly. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend prohibited-product content. [redacted] Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. The Advertising Policies prohibit advertising of vapes, vape products and other prohibited or regulated products outside permitted categories. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for the sale of prohibited and regulated products is conducted on a regular cadence; results are recorded in this Report and in Snap's Transparency Reports, which include a country-level breakdown of enforcement against illegal or regulated activities and dedicated breakdowns for drug-related and weapons-related violations. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA may concern content involving drug or weapons trafficking and are processed through the applicable priority-notice procedures. Snap also cooperates separately with specialist organisations and law-enforcement bodies, including the National Crime Agency in the UK in respect of the sale of knives and drugs. Further detail is provided in Section 5.10.

[redacted]

Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Cross-industry signal sharing is in operation, including the signal-sharing arrangement Snap operates with [redacted] in respect of illicit drugs content. Snap participates in awareness-raising initiatives, including a sustained partnership with the U.S. Ad Council on the Fentanyl Awareness Campaign and partnerships with Song for Charlie, the CDC, SAMHSA, Community Anti-Drug Coalitions of America (CADCA), Shatterproof, Truth Initiative and the SAFE Project.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Illegal or Regulated Activities category explainer, the Safety Center and Heads Up provide information on the substantive prohibitions and on resources available to users searching for drug-related terms. The Heads Up Search intervention, including the localised safety-support portals available during the assessment period, is recorded in Section 5.1.3.
Protection of Minors - Art. 35(1)(j)	Applies. Minor-protection measures specifically applicable to this risk include: (i) the minor-account contact-restriction defaults described in Section 5.1.3, which limit unsolicited contact from non-connected adult accounts [redacted]; (ii) Family Center oversight of teen-account contacts; (iii) [redacted] and (iv) [redacted] Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Partially applies. Content-authenticity measures (watermarking, contextual identification of generative-AI content) operate alongside the substantive prohibitions in the Community Guidelines and are described in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, assessed against an Article 35(1) proportionality frame, Snap has concluded that the mitigations in place for the dissemination of content related to the sale of prohibited or regulated products and services on Snapchat's in-scope services are reasonable, proportionate and effective.

This conclusion is reached on the strength of the broader mitigation set, including profile-level moderation, Search and Find Friends controls, the Proactive Safety Detection System and related enforcement measures. Heads Up operates as a supplementary informational intervention for

predefined drug-related Search queries. The combination of the platform architecture (no marketplace, recommended content in public feed is moderated, short default retention for public stories (unless saved to Public Profile), and pre-distribution moderation), the proactive moderation suite (ALD, multi-modal classifiers, Proactive Safety Detection System, profile-level moderation), Article 22 DSA trusted-flagger cooperation and separate law-enforcement cooperation, cross-industry signal-sharing arrangements, and awareness-raising partnerships supports the conclusion that the mitigation framework operates effectively against this risk. The low absolute PVP rates recorded across the relevant categories support this conclusion.

The mitigation framework will continue to be monitored and adjusted as the operational outputs of the changes mature, and as new evasion patterns or distribution vectors are identified through ongoing risk-detection processes (Section 6) and through cross-industry and law-enforcement engagement.

4.1.4 Dissemination of Terrorist Content

This Section 4.1.4 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of terrorist content on Snapchat's in-scope services. The risk is assessed in the context of, and in addition to, Snap's obligations under Regulation (EU) 2021/784 on addressing the dissemination of terrorist content online (the "TCO Regulation"), which establishes a separate but related legal framework applicable to terrorist content.

Harm Description

Terrorism, for the purposes of this assessment, refers to content that promotes or supports terrorism or other violent, criminal acts committed by individuals or groups to further ideological goals (whether political, religious, social, racial or environmental in nature). It includes content that:

- promotes or supports any terrorist organisation, including content that advances recruitment for terrorist organisations;
- promotes, incites, glorifies or facilitates terrorist activity or ideology, including in the form of propaganda, symbols, video or text glorifying mass violence;
- advocates or advances violent extremism or hate-group activity directed at the commission of violent acts.

A summary and explanation of "terrorism" for moderation purposes is provided in the Community Guidelines explainer on Hateful Content, Terrorism and Violent Extremism, and in the Snap Transparency Report Glossary.

Likelihood

Internal data

[redacted]

Prevalence Testing for Terrorist Content has consistently recorded extremely low rates on Snapchat's in-scope services[35]. Terrorist Content PVP was [redacted] in H1 2026 at the displayed level of precision, compared with [redacted] recorded in the 2025 Report. The H1 2026 result represents a decrease from the prior assessment cycle and confirms that Terrorist Content remains among the lowest-prevalence categories on Snapchat's in-scope services. The displayed H1 2026 figure should not be interpreted as indicating the absolute absence of violative content, as the result is rounded to the precision shown.

Where Terrorist Content was identified proactively or following a report, the median turnaround time from detection to final action was 1 minute during H2 2025, as reported in Snap's H2 2025 global Transparency Report.

Six-month average VVR data for Terrorist Content for EU users and EU minors during first half of 2026 for Spotlight, Public Story and Maps are set out below:

Audience	Spotlight VVR	Public Story VVR	Maps VVR
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. The Lens review taxonomy does not include a terrorism-specific rejection category. During H1 2026, Lens submissions were rejected under broader potentially relevant categories including Harassment/Hate Speech [redacted] promotion of dangerous, harmful or illegal activity [redacted] and content inciting or glorifying violence or the use of weapons [redacted] These figures are not limited to Terrorist Content and should not be interpreted as terrorism-specific rejection volumes.

External evidence

External evidence relevant to this risk includes:

- **EU TCO Regulation removal orders.** The European Commission reported in February 2024 that approximately 350 removal orders had been issued under the TCO Regulation since it began applying on 7 June 2022. Snap's records indicate that none of those orders was directed at Snap.[36]
- [redacted]

[35] Spotlight, Public Story, Maps, Saved Story, Communities.

[36] European Commission, 'Commission Reports on the Effective Implementation of the Terrorist Content Online Regulation' (14 February 2024).

[redacted]

- **Sector-wide studies.** Tech Against Terrorism's April 2023 *Patterns of Online Terrorist Exploitation* report found that file-sharing and archiving services were the platform types most at risk of terrorist exploitation and that messaging platforms hosted the majority of the far-right terrorist content identified in its dataset. Research published by the Counter Extremism Project, including its *Extremist Content Wrap Up 2022* and *Terrorists on Telegram*, likewise documents terrorist and extremist use of messaging, video, social-media and file-hosting services. These studies provide relevant sector-wide context but do not provide a Snapchat-specific prevalence measure.[37]
- **Internal media analysis.** Internal media analysis identified, in 2025, two specific UK-related incidents involving Snapchat references: a 2023 case in which an Essex teenager was convicted of preparing terrorist acts (in which Snapchat was used to post a single image as part of a broader course of conduct), and a 2022 incident involving a UK student making an inappropriate "joke" referring to terrorism (which did not result in terrorism charges). Neither incident is indicative of Snapchat being a primary distribution surface for terrorist content.

Likelihood conclusion

The internal Prevalence Testing data, including the H1 2026 PVP result of [redacted] at the displayed level of precision, the rapid median turnaround time for enforcement action, the very low VVR figures across the assessed in-scope services, including for EU minor users, the absence of TCO Regulation removal orders directed at Snap, and the consistent third-party intelligence-vendor and sector-wide-study findings together support the determination that the likelihood of dissemination of terrorist content on Snapchat's in-scope services is at the **lowest likelihood category**. There is no change in this determination from the 2025 Report. The decrease from the result recorded in the 2025 Report further supports the assessment that Terrorist Content remains among the lowest-prevalence categories on the in-scope services.

Severity

The European Commission has stated that the presence of terrorist content online represents a grave risk to citizens and to society at large. The Council of Europe, the Global Internet Forum to Counter Terrorism (GIFCT) and Member State counter-terrorism authorities have, on a sustained basis, recognised the harm associated with terrorist content distribution online - including the radicalisation of "lone wolves" responsible for terrorist acts in Europe, the targeting of young people for recruitment, and the enduring effects on victims and witnesses of attacks.

[37]Tech Against Terrorism, Patterns of Online Terrorist Exploitation: TCAP Insights (28 April 2023); Counter Extremism Project, Extremist Content Wrap Up 2022 (10 January 2023); Counter Extremism Project, Terrorists on Telegram.

[redacted]

The severity of negative effects associated with terrorist content is assessed as **serious harm industry-wide**, consistent with prior Reports. The classification reflects the gravity of harm associated with successful radicalisation, recruitment or attack-planning facilitated by online content, while distinguishing the category from those associated with imminent and direct threats to life on the same scale as severe-harm categories (for example, CSEA Content). As a general matter, and consistent with Section 3.5, any specific terrorism-related instance presenting an imminent, credible threat to life is treated as a Level 1 overall risk and handled accordingly through the Content Crisis Response Protocol described in Section 5.12.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services. The relevance of each factor to this systemic risk is considered below. The corresponding mitigations are assessed separately under Article 35.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend Terrorist Content; no corresponding interest category exists. The absence of an open recommender feed for unvetted user-generated content, the absence of a broad reshare functionality of Stories, [redacted] Further detail in Section 5.6 (Algorithmic Systems).
Content moderation systems	Proactive and reactive moderation procedures operate to identify and remove Terrorist Content. Automated detection includes ALD [redacted] multi-modal classifiers for hate-related and terrorism-related categories; and the Proactive Safety Detection System for [redacted] content at upload. The expansion of profile-level moderation in Section 5.4.4 supports the detection framework at the username and profile-pattern levels. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines prohibit the use of Snapchat by terrorist organisations, violent extremists and hate groups, and prohibit content that advocates or advances terrorism or violent extremism. The Community Guidelines apply uniformly across the in-scope services. Severe-harm enforcement consequences apply to terrorism-related violations. Further detail in Section 5.2 and Section 5.5.

[redacted]

Systems for selecting and presenting advertisements	The Advertising Policies prohibit advertising containing terrorist or violent-extremist content. Advertisements are subject to a mix of pre-publication review using automated screening and risk-based human review. Further detail in Section 5.7 (Advertising Systems).
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Where terrorism-related content is identified proactively or reactively, the median turnaround time from detection to final action was 1 minute during H2 2025, as reported in Snap's H2 2025 global Transparency Report.
Regional and linguistic aspects	ALD banned-word lists for terrorism-related content are maintained across multiple languages and categorised by context and region. Moderation guidance is calibrated to Member State legal contexts where required (for example, in respect of national-level terrorist designations and proscribed-organisation listings). The Content Crisis Response Protocol applies to elevated-risk events with potential Member State-specific dimensions.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, dissemination of terrorist content is assessed at **Level 2** overall potential risk prioritization. This prioritization reflects a case-by-case upward departure from the default Level 3 matrix output for serious harm industry-wide combined with lowest likelihood. The Level 2 prioritization reflects the potentially

[redacted]

grave public-security consequences of successful radicalisation, recruitment or attack-planning, and the need for continued heightened monitoring notwithstanding the very low observed prevalence. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to limit the structural opportunity for terrorist-content distribution: recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines prohibit the use of Snapchat by terrorist organisations, violent extremists and hate groups, and prohibit content that advocates or advances terrorism or violent extremism. Enforcement consequences apply at the severe-harm level, including, in the ordinary course, the locking of associated accounts and content retention to support law-enforcement investigation. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Proactive moderation includes ALD scanning [redacted] multi-modal classifiers; and the Proactive Safety Detection System with auto-enforcement [redacted] Reactive moderation is available through in-app reporting under the illegal content reporting option (Section 5.4.6) and the broader Community Guidelines reporting categories. Further detail in Section 5.4.

[redacted]

Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend Terrorist Content; no corresponding interest category exists. Pre-distribution moderation applies to content recommended on Spotlight and Discover. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising policies prohibit terrorist content. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Terrorist Content is conducted on a regular cadence; results are recorded in this Report and in Snap's Transparency Reports, which include EU-specific enforcement-volume data. [redacted]
Trusted Flaggers - Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA may concern Terrorist Content and are processed through the applicable priority-notice procedures. Orders issued by competent authorities under the TCO Regulation are processed separately through Snap's dedicated law-enforcement and TCO-response pathways. Orders issued under Articles 9 and 10 DSA are processed through the corresponding DSA order-handling pathway. Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap participates in the [redacted] and engages with cross-industry signal-sharing arrangements relevant to terrorism-related content. Snap operates an internal Content Crisis Response Protocol (Section 5.12) that applies in the event of major incidents that may give rise to an elevated terrorism-related risk.

[redacted]

Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Hateful Content, Terrorism and Violent Extremism category explainer (in the Community Guidelines Explainer Series), the Safety Center, and Snap's Transparency Reports provide information on terrorism-related policy, enforcement and prevalence. Statements of reasons under Article 17 DSA are issued in connection with enforcement decisions. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Mitigations for minors include the contact-restriction defaults described in Section 5.1.3, Family Center oversight (Section 5.8), and the elevated severity treatment applied to any terrorism-related content involving or directed at minor accounts. [redacted]
Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines prohibitions apply without regard to whether terrorism-related content is user-generated or AI-generated. Content-authenticity measures (watermarking, contextual identification of generative-AI content) operate alongside the substantive prohibitions and are described in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of terrorist content on Snapchat's in-scope services are reasonable, proportionate and effective.

Terrorist content remains one of the lowest-prevalence categories on Snapchat's in-scope services, and external evidence (including TCO Regulation removal-order data, [redacted] and sector-wide study findings) consistently indicates that Snapchat is not a primary distribution surface for this category. The combination of the platform architecture (recommended content in public feed is moderated; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation), the proactive moderation suite (ALD, multi-modal classifiers, Proactive Safety Detection System [redacted] profile-level moderation), the rapid median turnaround time for enforcement action, the cross-industry cooperation through the EU Internet Forum, and the Content Crisis Response Protocol for elevated-risk incidents, supports the conclusion that the mitigation framework operates effectively against this risk.

[redacted]

The mitigation framework will continue to be monitored on an ongoing basis. The H1 2026 PVP result of [redacted] at the displayed level of precision represents a decrease from the result recorded in the 2025 Report and supports the conclusion that Terrorist Content remains among the lowest-prevalence categories on Snapchat's in-scope services. Any material change in the prevalence result or broader risk profile, including in response to events that may give rise to an elevated terrorism-related threat profile, will be addressed through the Content Crisis Response Protocol and reflected in subsequent Reports as appropriate.

4.1.5 Dissemination of Content that Infringes on Intellectual Property Rights

This Section 4.1.5 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of content that infringes on the intellectual-property rights of others on Snapchat's in-scope services.

Harm Description

Intellectual-property infringement, for the purposes of this assessment, refers to the dissemination of content that, without authorisation, reproduces, distributes, communicates to the public, or otherwise infringes the rights of holders of copyrighted works, registered trademarks, or other intellectual-property rights protected under EU or Member State law. The category covers, in particular:

- the unauthorised reproduction or distribution of copyrighted material (including audiovisual works, music, photographs and other protected works);
- the use of registered trademarks in a manner that is likely to cause consumer confusion or that constitutes infringement under the applicable trademark framework;
- the use of intellectual property in connection with the offer or sale of counterfeit goods (which also falls within the scope of Section 4.1.3 in respect of counterfeit-goods sale);
- other infringements of intellectual-property rights protected under EU or Member State law.

The substantive prohibitions in respect of this category are set out in the Snap Terms of Service, the Community Guidelines, and Snap's Copyright and Trademark Policies. Reporting pathways are available to rights-holders through the in-app reporting menu, through the dedicated Copyright Infringement and Trademark Infringement online forms, and through the Article 16 DSA notice-and-action mechanism.

Likelihood

Internal data

[redacted]

Notice volumes and outcomes for IP-related enforcement on Snapchat's in-scope services[38] are recorded in Snap's Transparency Reports. The H2 2025 global Transparency Report disclosed the following volumes:

Category	Volume and outcome
Copyright notices (H2 2025)	1,213 notices received; 76.8% led to removal of some content
Copyright notices (H2 2024 for comparison)	1,296 notices received; 96.8% led to removal of some content
Trademark notices (H2 2025)	142 notices received; 55.5% led to removal of some content
Trademark notices (H2 2024 for comparison)	169 notices received; 60.9% led to removal of some content

These volumes reflect a low absolute incidence of IP-related notices on the in-scope services. Notice volumes decreased year over year, from 1,296 to 1,213 for copyright and from 169 to 142 for trademark. During H2 2025, 76.8% of copyright notices and 55.5% of trademark notices led to removal of some content.

No separate PVP or VVR is presented for this category because determining whether content infringes intellectual-property rights may depend on ownership, authorisation, licensing, applicable exceptions and information supplied by rights-holders that cannot reliably be determined from the content or view alone. The likelihood assessment therefore relies principally on rights-holder notice volumes and removal outcomes.

Advertising and Lenses are assessed through separate pre-publication review, rejection, post-publication reporting and enforcement datasets. During H1 2026, [redacted] Lens submissions were rejected for containing copyrighted or trademark content belonging to third parties.

External evidence

The European Union Intellectual Property Office (EUIPO) monitors trends in digital piracy and intellectual-property crime. EUIPO's 2024 study of online copyright infringement across the period 2017–2023 found that web-based piracy declined until late 2021 before increasing in 2022, while remaining significantly below 2017 levels. A joint EUIPO–Europol report published in

[38]Spotlight, Public Story, Maps, Saved Story, Communities

2024 identified social media and social commerce, including influencer promotion, as channels through which consumers may be directed towards counterfeit goods and pirated content.[39]

As recorded in the DSA Risk Factors below, Snapchat's platform architecture is meaningfully different from the platform types that the EUIPO and EUIPO-Europol reports principally focus on. In particular, the absence of a persistent, easily-searchable public content library, short default retention for public stories (unless saved to Public Profile), and the absence of a broad reshare functionality of stories reduce the structural opportunity for the large-scale dissemination of unauthorised copyrighted content.

Likelihood conclusion

On the basis of the internal Transparency Report notice volumes, the recorded removal outcomes following notice, and the platform-architecture factors described above, the likelihood of dissemination of content infringing on intellectual-property rights on Snapchat's in-scope services is assessed at the **lowest likelihood** category. There is no change in this determination from the 2025 Report.

Severity

Intellectual-property infringement does not, in the ordinary course, involve harms that risk significant damage to the physical or emotional well-being of users, nor does it ordinarily involve an imminent, credible risk of severe harm to human life, safety or well-being. The category is, however, the subject of substantial international, EU and national legal frameworks. Intellectual-property law confers, as a general matter, exclusive rights of exploitation on rights-holders and protects against unauthorised reproduction and distribution of protected works.

Where infringing content is disseminated at scale, the harm to rights-holders can be material in economic terms, and the broader harm to the integrity of the IP-rights framework can be material in policy terms. Recognising both the legally-protected nature of the category and the absence, in the ordinary course, of imminent harm to physical or emotional well-being, the severity of negative effects associated with this category is assessed as **significant harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services. The relevance of each factor

39 https://www.euipo.europa.eu/en/publications/online-copyright-infringement-in-the-european-union-films-music-publications-software-and-tv-2017-2023?.

[redacted]

to this systemic risk is considered below. The corresponding mitigations are assessed separately under Article 35.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend content that infringes intellectual-property rights; no corresponding interest categories exist. The absence of an open recommender feed for unvetted user-generated content, the absence of a broad reshare functionality of Stories, and pre-distribution moderation for content reaching wide audiences reduce the structural opportunity for the persistent and large-scale distribution of unauthorised copyrighted content. Further detail in Section 5.6.
Content moderation systems	[redacted] Statements of reasons under Article 17 DSA are issued in connection with enforcement decisions. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Snap Terms of Service, the Community Guidelines, and Snap's Copyright and Trademark Policies prohibit the dissemination of content that infringes intellectual-property rights. A repeat-infringer policy applies to copyright violations: where Snap becomes aware that a user has repeatedly infringed copyrights, account-level enforcement steps are taken consistent with the Terms. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising that infringes intellectual-property rights. Advertising review applies to all advertising prior to publication. Further detail in Section 5.7.
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework described in Section 6.3, including in respect of generative-AI features the outputs of which could[redacted] surface unauthorised reproductions of protected content.
Intentional manipulation (Article 34(2))	[redacted]

[redacted]

Amplification and rapid, wide dissemination	The platform architecture limits the structural opportunity for amplification of infringing content: short default retention for public stories (unless saved to Public Profile); the absence of an open recommender feed for unvetted user-generated content; the absence of a broad reshare functionality of Stories; and the pre-distribution moderation applied to content reaching wide audiences on Spotlight and Discover. The architectural focus on transient, curated sharing reduces the platform's utility for large-scale copyright infringement.
Regional and linguistic aspects	The Article 16 DSA notice-and-action mechanism, the dedicated Copyright and Trademark reporting forms, and the in-app reporting flow are available in multiple EU languages. Member State-specific copyright and trademark frameworks are reflected in the substantive Terms and the moderation framework where required.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, dissemination of content infringing on intellectual-property rights is assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of significant harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

[redacted]

Snapchat Design and Function - Art. 35(1)(a)	Applies. The platform architecture limits the structural opportunity for IP infringement: short default retention for public stories (unless saved to Public Profile); the absence of an open recommender feed for unvetted user-generated content; the absence of a broad reshare functionality of Stories; pre-distribution moderation for content reaching wide audiences. Reporting tools support the detection of infringing content. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Terms of Service, Community Guidelines, and Snap's Copyright and Trademark Policies prohibit IP infringement. Where a rights-holder believes that content on Snapchat infringes their intellectual-property rights, they may submit a report through the in-app reporting menu, through the dedicated Copyright Infringement and Trademark Infringement online forms, or through the Article 16 DSA notice-and-action mechanism. Snap takes steps to expeditiously remove infringing material upon becoming aware of it. The repeat-infringer policy applies to copyright violations. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Specific [redacted] moderation procedures operate to expeditiously remove content that infringes intellectual-property rights upon receipt of a notice. [redacted] Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend content that infringes intellectual-property rights; no corresponding interest categories exist. Pre-distribution moderation applies to content recommended on Spotlight and Discover. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit advertising that infringes intellectual-property rights. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. IP-related notice volumes and removal outcomes are monitored on a continuous basis through the Content Moderation Health Monitoring Dashboard, and are disclosed in Snap's Transparency Reports. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies, in limited form. Notices submitted by entities with established rights-holder status are processed through the dedicated Copyright and Trademark reporting forms, which provide for expedited handling.

Codes and Crisis Protocols - Art. 35(1)(h)	Limited application. No specific code of conduct is in operation in respect of IP infringement on the in-scope services. Snap engages with rights-holder organisations and industry bodies on an ad hoc basis. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. Snap's Transparency Reports include data on enforcement actions related to IP infringement, including copyright and trademark notice volumes and removal outcomes, with a country-level breakdown for EU Member States. Statements of reasons under Article 17 DSA are issued in connection with IP-related enforcement decisions. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies, in limited form. The general minor-protection measures described in Section 5.8 apply. IP-infringement risk does not, in the ordinary course, present a minor-specific harm vector that warrants additional minor-specific mitigation beyond the general framework.
Content Authenticity - Art. 35(1)(k)	Partially applies. Generative-AI features made available on Snapchat operate within content-policy and rights-handling frameworks designed to mitigate the risk that their outputs could surface unauthorised reproductions of protected content. Watermarking and contextual identification of generative-AI content are described in Section 5.9 and operate alongside the substantive prohibitions applicable to IP infringement.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of content infringing on intellectual-property rights on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (short default retention for public stories (unless saved to Public Profile), recommended content in public feed is moderated, no broad reshare functionality of Stories), the dedicated Copyright and Trademark reporting forms operating alongside the Article 16 DSA notice-and-action mechanism, the recorded removal outcomes following notice, the repeat-infringer policy, and Snap's Transparency Report disclosure of category-specific notice volumes and outcomes, supports the conclusion that the mitigation framework operates effectively against this risk. There is no change in the overall conclusion from the 2025 Report.

4.1.6 Dissemination of Adult Sexual Content

This Section 4.1.6 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of Adult Sexual Content on Snapchat's in-scope services. The risk is sub-divided for severity-classification purposes into two categories: sexual crimes (Serious industry-wide) and other Adult Sexual Content (Significant industry-wide), each assessed at lowest likelihood. This section addresses Adult Sexual Content; content concerning child sexual abuse material is addressed in Section 4.1.1 (CSAM).

Harm Description

Adult Sexual Content, for the purposes of this assessment, encompasses pornographic content as well as commercial activities that relate to pornography or sexual interactions (whether online or offline), including the sending of unsolicited explicit images or videos. Breastfeeding and other depictions of adult nudity in non-sexual contexts are, as a general matter, permitted under the Community Guidelines. The detailed substantive prohibitions are recorded in the Community Guidelines on Sexual Content, the corresponding Community Guidelines Explainer, and the Snap Transparency Report Glossary.

Sub-categories

Sexual crimes

This sub-category covers categories of Adult Sexual Content that are or may be illegal, including:

- Sexual exploitation, including the criminal exploitation of vulnerable persons through sexual means;
- Sexual harassment, including unwanted advances, sharing of graphic and unsolicited content, and obscene requests;
- Sextortion (including financial sextortion), being the use of intimate imagery as a tool for coercion or extortion;
- Non-consensual intimate imagery (NCII), including the production, sharing or threatened sharing of intimate imagery without the depicted person's consent ("revenge pornography");
- Deceptive sexual practices, including the use of deception to coerce or entice users to provide intimate content;
- Commercial sex transactions, where these constitute or facilitate offences under applicable law (including in connection with sex trafficking).

Other Adult Sexual Content

This sub-category covers pornographic content, sexual nudity where the primary intention is sexual arousal, and other depictions of explicit sex acts, that fall within the scope of the Community Guidelines prohibitions but do not, in the ordinary course, constitute illegal conduct.

[redacted]

Sextortion

Sextortion, including financial sextortion, is treated within this assessment as part of the sexual-crimes sub-category for severity-classification purposes. Where sextortion involves contact with, or solicitation of, minors, the associated risk pathways are addressed across Sections 4.1.1, 4.1.6, 4.1.7, 4.4.2 and 4.4.3, with related safeguards described in Sections 5.4 and 5.8. The mitigations described here, including in-app warnings, friending-protection measures, in-app educational resources, the dedicated reporting reason for financial sextortion, NCMEC's Take It Down service for imagery involving minors, and SWGfL's StopNCII service for adult intimate imagery, operate alongside the broader minor-protection framework described in Section 5.8 and Section 6.1.

Likelihood

Internal data

Prevalence Testing for Adult Sexual Content on Snapchat's in-scope services[40] has recorded a substantial long-term decline across multiple assessment cycles, with prevalence remaining broadly stable at a low level in the 90-day snapshot ending 30 June 2026:

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Using the methodology retained for year-on-year comparability in Section 6.4, Adult Sexual Content PVP was [redacted] for the 90-day period ending 30 June 2026, compared with [redacted] recorded in the 2025 Report. This represents a modest decrease under the comparable initial-review methodology, with prevalence remaining substantially below the levels recorded in earlier assessment cycles. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP for the 90-day period ending 30 June 2026 was [redacted] The refined figure is presented as supplementary 2026 context and is not used to

[40] Spotlight, Public Story, Maps, Saved Story, Communities.

[redacted]

determine the year-on-year direction. Adult Sexual Content remained the highest-prevalence harm category identified through Snap's Prevalence Testing.

The H2 2025 global Transparency Report recorded a median turnaround time of 1 minute, whether identified proactively or following a report.

Violative View Rate (H1 2026)

Six-month average VVR data for Adult Sexual Content sub-categories for EU users during the period from 1 January to 30 June 2026 are set out below for Spotlight, Public Story and Maps.

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

The corresponding VVR data for EU minor users aged 13–17 were as follows:

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

The VVR figures above are disaggregated by individual enforcement reason and public surface and are therefore considered alongside the aggregate PVP measure for the Adult Sexual

[redacted]

Content category. In particular, the Public Story VVR for Sexual Nudity/Pornography and Commercial Sex Transactions is materially higher than the corresponding VVR on Spotlight and Maps and is treated as a surface-specific signal requiring continued monitoring. In accordance with the methodology in Section 3.2, Snap has considered these VVR results together with the aggregate PVP result, the VVR results across the other public surfaces and for minor users, enforcement turnaround times, and the platform-architecture factors described below. On the totality of that evidence, Snap maintains the lowest-likelihood classification for both Adult Sexual Content sub-categories.

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. For Lenses, during H1 2026, [redacted] submissions were rejected under the operational category "Includes depictions of nudity or sex acts." This category is broader than the Adult Sexual Content category assessed for this risk and should not be interpreted as a PVP or VVR measure.

External evidence

External evidence relevant to this risk includes:

- **EU and UN findings on sexual violence and exploitation.** The 2026 EU Gender-Based Violence Survey published by FRA and EIGE found that 30.7% of women in the EU had experienced physical and/or sexual violence and 17.2% had experienced sexual violence. The survey also documents physical injuries, depression, anxiety and other significant consequences for victims. UNODC's 2024 Global Report on Trafficking in Persons similarly identifies sexual exploitation as a major form of trafficking and documents the exploitation of people in vulnerable situations.[41]
- **Minor exposure and synthetic sexual imagery.** The European Commission has treated minors' exposure to pornographic content as a risk to their rights and wellbeing requiring effective age-assurance and mitigation measures under the DSA. Internet Matters has reported that the increased accessibility of generative-AI tools has made the production of realistic nude and sexual deepfakes substantially easier, creating risks of harassment, sexual abuse and sextortion involving children.[42]
- **Human trafficking and online facilitation.** Eurostat recorded 9,678 registered victims of trafficking in human beings in the EU in 2024. Among victims for whom the form of exploitation was reported, 4,219 (or 46%) were trafficked for sexual exploitation. Europol

[41] European Union Agency for Fundamental Rights (FRA), European Institute for Gender Equality (EIGE) and Eurostat, EU Gender-Based Violence Survey - Evidence for Policy and Practice (2026); United Nations Office on Drugs and Crime, Global Report on Trafficking in Persons 2024 (2024).

[42] European Commission, "Commission preliminarily finds PornHub, Stripchat, XNXX and XVideos in breach of the Digital Services Act for allowing minors to access their services," Press Release, 26 March 2026; Internet Matters, The New Face of Digital Abuse: Children's Experiences of Nude Deepfakes (October 2024).

has separately identified the increasing use of online services to recruit victims, build manipulative relationships and facilitate or livestream sexual exploitation.[43]

- **Cross-platform sextortion research.** Thorn's 2024 research identified social-media and messaging services as recurring components of financial-sextortion schemes and found that perpetrators frequently move victims to secondary, including end-to-end encrypted, messaging services. In separate 2022 research, Thorn reported that 65% of children surveyed had experienced an online-only contact attempting to move them from a public chat to a private conversation on another platform.[44]

Likelihood conclusion

On the basis of the internal Prevalence Testing data, the long-term decline in PVP across multiple assessment cycles, the VVR data for EU users and EU minor users, the rapid median turnaround time for enforcement action, and the platform-architecture factors recorded above, the likelihood of dissemination of Adult Sexual Content on Snapchat's in-scope services is assessed at the lowest likelihood category for both sub-categories: sexual crimes and other Adult Sexual Content.

The 2025 Report recorded a movement of this category from medium to lowest likelihood. Using the methodology retained for year-on-year comparability, the 2026 PVP result of [redacted] for the 90-day period ending 30 June 2026 was broadly stable relative to the [redacted] recorded in the 2025 Report and remained substantially below earlier assessment-cycle levels. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP was [redacted] Although certain enforcement-reason and surface-specific VVR figures fall within higher quantitative bands under Section 3.2, those figures are assessed as part of the totality of the evidence for the reasons set out above. On that basis, the evidence supports maintaining the lowest relative likelihood classification for both sub-categories in 2026.

Severity

Sexual crimes

Sexual offences – including sexual exploitation, sexual harassment, sextortion, NCII, and sex trafficking-adjacent activity -- pose a significant threat to human life, dignity and well-being, as recognised by the EU Charter, the European Convention on Human Rights, and the corresponding international instruments. Recent data shows technology-mediated gender-based violence has been recognised as a serious global-scale problem, exacerbated by the increasing

[43] Eurostat, "9 678 registered victims of human trafficking in 2024," 29 January 2026; Europol, "Over 30 potential victims identified in action against human trafficking enabled online," 13 October 2025.

[44] Thorn, Trends in Financial Sextortion: An Investigation of Sextortion Reports in NCMEC CyberTipline Data (June 2024); Thorn, Online Grooming: Examining Risky Encounters Amid Everyday Digital Socialization (2022).

[redacted]

availability of generative-AI tools. The severity classification for sexual crimes is **serious harm industry-wide**.

Other Adult Sexual Content

The dissemination of pornographic content, sexual nudity where the primary intention is sexual arousal, and other depictions of explicit sex acts, is associated with significant harm industry-wide. The Council of Europe's Parliamentary Assembly has noted that pornography can engender and perpetuate stereotypes that undermine gender equality and women's self-determination. Research has noted associations between online pornography consumption and certain forms of physical or verbal violence against women. Particular concern arises in respect of minor exposure, including (as noted by the European Commission) where exposure at early ages may affect views of healthy relationships. The severity classification for other Adult Sexual Content is **significant harm industry-wide**.

As a general matter, and consistent with Section 3.5, any specific instance within either sub-category that presents an imminent, credible risk of severe harm - including a threat to human life, safety or well-being (for example, in connection with active sextortion targeting a minor or active sex-trafficking activity) - is treated as a Level 1 overall risk and handled accordingly.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend Adult Sexual Content; no corresponding interest category exists. The platform is designed to limit the prevalence of sexually-suggestive content reaching wide audiences. The absence of an open recommender feed for unvetted user-generated content, the absence of a broad reshare functionality of Stories, and pre-distribution moderation for content reaching wide audiences reduce the structural opportunity for amplification. Further detail in Section 5.6.
Content moderation systems	Proactive and reactive moderation procedures operate to detect and remove Adult Sexual Content. Automated detection [redacted] are used for sexual-exploitation imagery

[redacted]

	(including in support of CSAM detection per Section 4.1.1). Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Sexual Content prohibit (i) sexual exploitation, sexual harassment, sextortion, deceptive sexual practices and commercial sex transactions; (ii) the production, sharing or threatened sharing of non-consensual intimate imagery (NCII); and (iii) the promotion, distribution or sharing of pornographic content, sexual nudity, or offers of sexual services. Severe-harm enforcement consequences apply to sexual-crime categories, including account-level enforcement and content retention to support law-enforcement investigation. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit Adult Sexual Content in advertisements. Advertising review applies to all advertising prior to publication. Further detail in Section 5.7.
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3, including in respect of generative-AI features the outputs of which could, in the absence of mitigations, surface sexually-explicit content (including synthetic NCII).
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of Adult Sexual Content on in-scope services is constrained by the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation) and by the auto-enforcement threshold applied at upload. Where Adult Sexual Content is identified proactively or reactively, the median turnaround time from detection to final action was 1 minute during H2 2025, as reported in Snap's H2 2025 global Transparency Report.

[redacted]

Regional and linguistic aspects	ALD banned-word lists for sexual content are maintained across multiple languages and categorised by context and region. Trusted-flagger relationships include Member State-specific organisations (notably Stop Fisha in France in respect of NCII). Moderation guidance is calibrated to Member State legal contexts where required.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, the overall risk prioritisation for this risk is recorded by sub-category as follows:

[redacted]

The matrix-level positioning of each sub-category is shown below:

[redacted]

The determination for Other Adult Sexual Content corresponds to the matrix default for significant harm industry-wide combined with lowest likelihood and is therefore Level 3. For sexual crimes, the matrix default for serious harm industry-wide combined with lowest likelihood is Level 3; however, applying the case-by-case deviation framework in Section 3.5, sexual crimes are assessed at Level 2 overall potential risk prioritisation. This upward adjustment reflects the potential seriousness of sexual exploitation, sextortion, NCII and related sexual offences, including risks to vulnerable users and the potential for serious and lasting harm where such conduct materialises. There is no change in either determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to limit the structural opportunity for Adult Sexual Content distribution: recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences. Minor-account default settings (Section 5.1.3) further limit minor exposure to Adult Sexual Content. Friending-protection measures, including default contact restrictions and enhanced safeguards against suspicious friend requests, reduce the contact pathway for sextortion. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Sexual Content prohibit the categories described in the Harm Description above. Enforcement applies severe-harm consequences for sexual-crime categories, including account-level enforcement and, in the ordinary course, content retention to support law-enforcement investigation. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Proactive moderation includes [redacted] Reactive moderation is available through in-app reporting, including a tailored reporting reason for financial sextortion (introduced in 2024). [redacted] are deployed for known and novel sexual-exploitation imagery. NCMEC's Take It Down and SWGfL's StopNCII technology are integrated to support the prevention of underage and adult NCII spread. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend Adult Sexual Content; no corresponding interest category exists. Pre-distribution moderation applies to content recommended on Spotlight and Discover, including filtering of sexually-suggestive content from minor-user content feeds. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit Adult Sexual Content. Dating-services advertising is restricted to 18+ audiences [redacted]

[redacted]

	[redacted] Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Adult Sexual Content is conducted on a regular cadence; results are recorded in this Report and in Snap's Transparency Reports, with sub-category breakdown. The 2025 Report recorded a substantial long-term decline in PVP across multiple assessment cycles. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Trusted-flagger cooperation relevant to NCII and sexual exploitation includes Snap's relationship with Stop Fisha (France) in respect of NCII. Notices submitted by designated trusted flaggers under Article 22 DSA are separately subject to the priority-treatment framework described in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap participates in the [redacted] which has expanded its remit to cover human trafficking (which is, in many cases, driven by sexual crimes or pornography). Snap has invested in cross-platform research on sextortion and has engaged with industry peers on signal sharing for sexual-exploitation patterns. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Sexual Content category explainer (in the Community Guidelines Explainer Series), the Safety Center, and Snap's Transparency Reports provide information on the substantive prohibitions, on enforcement and prevalence, and on the resources available to users (including sextortion educational resources). In-app educational resources for teens regarding the threat of financial sextortion have been deployed. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Mitigations specific to minors include: (i) the contact-restriction defaults (Section 5.1.3); (ii) the enhanced friending-protection measures [redacted] (iii) [redacted] (iv) the filtering of sexually-suggestive content from minor-user content feeds; (v) the exclusion of certain minor-account contexts from suggestion to non-connected adult accounts; and (vi) Family Center oversight (Section 5.8). Further detail in Section 5.8.

[redacted]

Content Authenticity - Art. 35(1)(k)	Applies. The Community Guidelines apply without regard to whether Adult Sexual Content is user-generated or AI-generated, including in respect of synthetic NCII (deepfake nudes). Generative-AI Lens flows include input/output scanning using internal and external models, retry logic for unsafe outputs, and watermarking of generative-AI outputs. The 48-hour input-image deletion for high-latency Lens generation reduces the misuse pathway. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of Adult Sexual Content on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation, private friend lists, bidirectional friending requirement), the proactive moderation suite [redacted], the Proactive Safety Detection System, [redacted] the friending-protection measures, trusted-flagger cooperation, including with Stop Fisha, and the Article 22 DSA priority-treatment framework, and the dedicated minor-protection framework support the conclusion that the mitigation framework operates effectively against this risk. The substantial long-term decline in PVP across multiple assessment cycles supports this conclusion.

Adult Sexual Content remains the highest-prevalence violating category in absolute terms among the categories assessed in this Report, and the mitigation framework will continue to be subject to careful ongoing monitoring with the aim of achieving further prevalence reductions. There is no change in the overall conclusion from the 2025 Report.

4.1.7 Dissemination of Content regarding Harassment and Bullying

This Section 4.1.7 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of content constituting harassment, bullying or stalking on Snapchat's in-scope services.

Harm Description

Snap defines harassment as any unwanted behaviour that could cause an ordinary person to experience emotional distress, including bullying, verbal abuse and certain categories of threatening behaviour. Examples of conduct that fall within this category include:

[redacted]

- aggressive or profane name-calling;
- shaming or embarrassing imagery directed at a person's physical appearance, disabilities or cognitive abilities;
- the use of a new account to contact a person after the user has previously been blocked by them;
- attempts at intimidation;
- "gossip" accounts whose primary purpose is to anonymously bully or shame individuals known to the account's followers;
- content that negatively focuses on a named or visible person's physical appearance, personal traits, cognitive or physical abilities, or economic circumstances;
- the sharing of images of a person in a private space (such as a bathroom, bedroom, locker room or medical facility) without that person's knowledge and consent;
- the sharing of another person's private information without their knowledge and consent or for the purpose of harassment ("doxxing").

Shaming behaviour is treated as a form of bullying, and includes any behaviour intended to embarrass or humiliate the target, or to make the target feel as though they have done something improper or are unworthy of acceptance and belonging. The substantive prohibitions are recorded in the Community Guidelines on Harassment and Bullying, the corresponding Community Guidelines Explainer, and the Snap Transparency Report Glossary.

The category includes harassment that constitutes a criminal offence under EU or Member State law (in particular, certain categories of cyber-harassment, stalking and threats), as well as conduct that is incompatible with the Community Guidelines but does not necessarily constitute a criminal offence. The overwhelming majority of cases that fall within this category are not, in the ordinary course, illegal under applicable law, but do violate the Community Guidelines.

Likelihood

Internal data

Prevalence Testing for Harassment and Bullying on Snapchat's in-scope services has recorded a sustained downward trajectory across recent assessment cycles. Prevalence peaked at [redacted] in August 2023 and subsequently declined to [redacted] in July 2024, [redacted] in April 2025 and [redacted] for the 90-day period ending 30 June 2026, using the methodology retained for year-on-year comparability in Section 6.4:

Period	Policy Violating Prevalence (PVP)
[redacted]	[redacted]
[redacted]	[redacted]

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]

The 2026 PVP result of [redacted] represents a substantial further decline from the [redacted] result recorded as at 30 April 2025 and confirms that Harassment and Bullying remains within the lowest likelihood band under the methodology in Section 3.2. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP for the 90-day period ending 30 June 2026 was [redacted]. The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction.

The H2 2025 global Transparency Report recorded a median turnaround time of 1 minute for enforcement action against Harassment and Bullying content.

Distribution of reports across public and private surfaces

For Spotlight, Discover, Public Profiles, and Snap Map, internal data for H2 2025 and H1 2026 shows the following distribution of Harassment-related reports submitted by minor users between private, out-of-scope surfaces and public, in-scope surfaces:

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

Although the overall volume of reports increased between the two periods, the distribution remained broadly stable, with approximately [redacted] of Harassment-related reports submitted by minor users originating on private surfaces and approximately [redacted] originating on public surfaces. This structural concentration in private contexts is consistent with Snapchat's principal use case of private communication between bidirectional friends and indicates that the in-scope public-surface risk is materially lower than the platform-wide reporting volume might otherwise suggest.

Violative View Rate (H1 2026)

Six-month average VVR data for Harassment and Bullying for EU users and EU minors during the period from 1 January to 30 June 2026 are set out below for Spotlight, Public Story and Maps:

Audience	Spotlight VVR	Public Story VVR	Maps VVR

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. For Lenses, during H1 2026, submissions were rejected under broader operational categories potentially relevant to harassment and bullying, including Harassment/Hate Speech [redacted] and targeted bullying or potential harassment of an individual [redacted] These categories are not limited to the Harassment and Bullying content assessed in this section and should not be interpreted as PVP or VVR measures.

Likelihood conclusion

On the basis of the internal Prevalence Testing data, the sustained downward PVP trajectory, the public-surface concentration analysis, the very low VVR figures (including for EU minor users), and the rapid median turnaround time for enforcement action, the likelihood of dissemination of Harassment and Bullying content on Snapchat's in-scope services is assessed at the **lowest likelihood** category. Similar to the 2025 Report, this year's determination preserves the lowest-likelihood placement in light of the [redacted] PVP result under the methodology retained for year-on-year comparability, the supplementary confirmed-violation PVP of [redacted] under the refined double-review methodology described in Section 6.4, the very low H1 2026 VVR figures and rapid median turnaround time, considered together with the public-surface concentration data.

Severity

Harassment and Bullying are recognised as serious threats to individual well-being and, in the EU regulatory context, are addressed by Member State criminal frameworks (in particular, in respect of cyber-harassment, stalking and threats) and by emerging EU-level instruments (including Directive 2024/1385/EU on combating violence against women, which addresses cyber-harassment as a category).

Specific risk pathways include: (i) emotional and psychological impacts on the targets of harassment, with particular concern in respect of vulnerable users including minors and members of marginalised communities; (ii) the potential for harassing conduct to escalate into physical-safety threats, including in connection with stalking and intimate-partner contexts; and (iii) the chilling effect of harassment on participation in online discourse, particularly for women and minority users.

Harassment and Bullying can cause serious emotional and psychological harm, particularly where conduct is sustained, targeted at vulnerable users, or occurs in stalking or intimate-partner

[redacted]

contexts. Although the category does not, in the ordinary course, present an imminent, credible threat to human life or severe physical safety, certain instances may escalate into physical-safety threats. The general severity classification for Harassment and Bullying is therefore assessed as serious harm industry-wide. Where a specific instance presents an imminent, credible threat to life or serious bodily injury, the Section 3.5 case-by-case framework applies and the matter is treated as a Level 1 overall risk through the prioritised-response and law-enforcement-escalation pathways.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend Harassment and Bullying content; no corresponding interest category exists. The absence of an open recommender feed for unvetted user-generated content, the absence of a broad reshare functionality of Stories, and pre-distribution moderation for content reaching wide audiences reduce the structural opportunity for amplification of harassing content. Further detail in Section 5.6.
Content moderation systems	Proactive and reactive moderation procedures operate to detect and remove Harassment and Bullying content. [redacted] human review applies to nuanced or borderline cases. Content can be reported in-app or through the Snapchat Support Site under the dedicated Harassment reporting reason. Account-level signals include tracking patterns of repeat reporting against the same target. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Harassment and Bullying prohibit the categories described in the Harm Description. Snap retains the ability to escalate to law enforcement where content suggests an emergency situation involving imminent danger of death or serious bodily injury; in the ordinary course, escalations to law enforcement for this category are rare, [redacted] Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising containing Harassment and Bullying content. Pre-publication review applies to all advertising. Further detail in Section 5.7.

[redacted]

Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of Harassment and Bullying content on in-scope services is constrained by the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation). Where Harassment and Bullying content is identified (proactively or reactively), the median turnaround time for enforcement action is rapid; the H2 2025 global Transparency Report disclosed a median turnaround time of 1 minute.
Regional and linguistic aspects	ALD banned-word lists for harassing content are maintained across multiple languages and categorised by context and region. Moderation guidance is calibrated to Member State legal contexts where required (for example, in respect of categories of cyber-harassment that are subject to specific Member State criminal legislation). The dedicated Harassment reporting reason is available in multiple EU languages.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, dissemination of Harassment and Bullying content is assessed at Level 2 overall potential risk prioritization. This prioritization reflects a case-by-case upward departure from the default Level 3 matrix output for serious harm industry-wide combined with lowest likelihood. The Level 2 prioritization reflects the potential for sustained or targeted harassment to cause serious psychological harm, particularly to minors and other vulnerable users, and the possibility of escalation into stalking or physical-safety threats. There is no change in this determination from the 2025 Report.

[redacted]

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)-(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. Foundational design choices reduce the structural opportunity for harassing content to reach a wide audience: recommended content in public feed is moderated; pre-distribution moderation for content reaching wide audiences; the absence of a broad reshare functionality of Stories. The block functionality, the simple in-app reporting flow, and the option to remove an account from contact (for friended accounts) provide users with direct controls over unwanted contact. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Harassment and Bullying prohibit unwanted behavior that could cause an ordinary person to experience emotional distress, including specific categories enumerated in the Harm Description above. Where content suggests an emergency situation involving imminent danger of death or serious bodily injury, response is prioritised, with escalation to law enforcement available where appropriate. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Specific moderation procedures operate for Harassment and Bullying content, [redacted] The simple in-app reporting flow includes a dedicated Harassment reporting reason; reporting is also available through the Snapchat Support Site. Group Chat Reporting (introduced in 2023) and Account Reporting Enhancements (introduced in 2024) allow users to provide chat evidence and to report multi-person chat contexts, expanding the reporting surface. Further detail in Section 5.4.

[redacted]

Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend Harassment and Bullying content; no corresponding interest category exists. Pre-distribution moderation applies to content recommended on Spotlight and Discover. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit advertising containing Harassment and Bullying content. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Harassment and Bullying is conducted on a regular cadence; results are recorded in this Report and in Snap's Transparency Reports. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies, in limited form. Bad-faith reporting of non-violating content as harassment is addressed through the Article 23 DSA misuse-suspension framework described in Section 5.4.6.
Codes and Crisis Protocols - Art. 35(1)(h)	Limited application. No specific code of conduct is in operation in respect of Harassment and Bullying on the in-scope services. Snap engages with industry peers and civil-society organisations on cross-platform harassment patterns where relevant. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Harassment and Bullying category explainer (in the Community Guidelines Explainer Series), the Safety Center, and Snap's Transparency Reports provide information on the substantive prohibitions, on enforcement and prevalence, and on the resources available to users. The Here for You in-app resource portal surfaces support content for users encountering harassment-related distress. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Mitigations specific to minors include: (i) the contact-restriction defaults (Section 5.1.3); (ii) Family Center oversight, including visibility of who teens are communicating with (Section 5.8); (iii) friending-protection measures, [redacted] (Section 5.3.3); and (iv) the in-app educational resources of Friend Check Up and Here for You. Further detail in Section 5.8.

[redacted]

Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines apply without regard to whether harassing content is user-generated or AI-generated. Generative-AI-created harassing content (including synthetic imagery used as a vehicle for harassment) is addressed under the substantive prohibitions, supported by the content-authenticity measures described in Section 5.9 (watermarking and contextual identification of generative-AI content).

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of Harassment and Bullying content on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (recommended content in public feed is moderated, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation, bidirectional friending), the proactive moderation suite [redacted], the simple in-app reporting flow with a dedicated Harassment reporting reason, the Group Chat Reporting and Account Reporting Enhancements features, the Family Center oversight for minor users, and the Here for You in-app resource portal supports the conclusion that the mitigation framework operates effectively against this risk. The very low H1 2026 VVR figures and rapid median turnaround time, considered together with the available historical prevalence and public-surface concentration data, support this conclusion.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes.

4.1.8 Dissemination of Content that Glorifies Self-Harm, including the Promotion of Self-Injury, Suicide or Eating Disorders

This Section 4.1.8 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of content that glorifies self-harm - including content promoting self-injury, suicide or eating disorders - on Snapchat's in-scope services. The category is one of those in respect of which Snap applies the heightened-attention framework for severe and serious harms recorded in Section 3 (Methodology).

Harm Description

[redacted]

This category, for the purposes of this assessment, refers to content relating to self-harm, suicide and eating disorders that meets the substantive prohibitions in the Community Guidelines on Threats, Violence and Harm. The category covers, in particular:

- content depicting self-harm or suicide posted for shock value rather than in a context of seeking or offering support;
- content promoting suicide or eating disorders by glamorising, romanticising or normalising the conduct;
- content that encourages, instructs or asks users to harm themselves;
- non-graphic statements towards self-harm, suicidal ideation, depression or other cause for concern, where the statements rise to the level of glorification or incitement;
- content that praises or glorifies eating disorders, including pro-restrictive or pro-purging content.

The substantive prohibitions are recorded in the Community Guidelines on Threats, Violence and Harm[45], the corresponding Community Guidelines Explainer[46], and the Snap Transparency Report Glossary[47]. The category does not, as a general matter, prohibit non-glorifying discussion of self-harm, suicide or eating disorders (for example, in news reporting, public-issue commentary, or peer-support contexts), as further described in the calibrated enforcement framework below.

Calibrated three-tier enforcement framework

Snap applies a calibrated enforcement framework to self-harm-related content, distinguishing three categories of content that warrant different response approaches:

[redacted]

[45] https://values.snap.com/policy/policy-community-guidelines
[46] https://values.snap.com/policy/policy-community-guidelines/threats-violence-harm
[47] https://values.snap.com/privacy/transparency/glossary

[redacted]

[redacted]

[redacted] It reflects the potentially complex role of self-harm-related content, including the importance of preserving access to legitimate help-seeking and peer-support content while preventing the dissemination and algorithmic amplification of harmful material.

Likelihood

Internal data

Using the methodology retained for year-on-year comparability in Section 6.4, Self-Harm and Suicide-specific PVP was [redacted] on Snapchat's in-scope services for the 90-day period ending 30 June 2026, compared with [redacted] recorded in the 2025 Report. This represents an increase under the comparable initial-review methodology, while prevalence remained very low in absolute terms. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP for the 90-day period ending 30 June 2026 was [redacted]. The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction.

The six-month average VVR data for Self-Harm and Suicide content for EU users and minors across the in-scope public surfaces of Spotlight, Public Stories and Maps during the period from 1 January to 30 June 2026 are set out below:

[redacted]

[redacted]

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. For Lenses, during H1 2026, [redacted] submissions were rejected under the operational category “Promotes harmful activity – self-harm.” This figure is a Lens-review rejection volume and should not be interpreted as a PVP or VVR measure.

Snap’s H2 2025 global Transparency Report recorded 39,437 enforcement actions against Self-Harm and Suicide content, affecting 34,950 unique accounts, with a median turnaround time of one minute from detection to final enforcement action.

External evidence

The risk of young people encountering content that glorifies self-harm, suicide or eating disorders on social-media services has been the subject of sustained external research and policy attention:

- **European Parliament research.** A 2023 study requested by the European Parliament’s Committee on Culture and Education examined the risks associated with children’s and young people’s use of social media, including exposure to harmful content and potential impacts on mental health and well-being. The study emphasised the need for children to be protected, safe and empowered when using online services.[48]
- **Reset research.** Research by Reset.Tech Australia tested the recommendation and moderation systems of several major platforms using accounts presented as children. The research identified the recommendation of pro-eating-disorder content on Instagram and X, as well as explicit self-harm imagery on X, illustrating the potential for recommender systems to expose young users to harmful material.[49]
- **Samaritans research.** Research by Samaritans and Swansea University found that 83% of surveyed social-media users had been recommended self-harm content on personalised feeds without searching for it. More recent Samaritans research has continued to examine how users encounter and engage with self-harm and suicide content online, including the importance of preventing harmful content from being algorithmically promoted while preserving access to appropriate support content.[50]
- **Molly Rose Foundation.** Research published by the Molly Rose Foundation in June 2026 surveyed 1,825 children aged 13–17 across 21 UK schools regarding their exposure to suicide, self-harm, depression and eating-disorder content on six major platforms, including Snapchat. Snapchat users were the least likely among the surveyed

[48] European Parliament, Directorate-General for Internal Policies, The Influence of Social Media on the Development of Children and Young People, study requested by the Committee on Culture and Education, February 2023.
[49] Reset.Tech Australia, Not Just Algorithms: Assuring User Safety Online with Systemic Regulatory Frameworks, March 2024, pp. 9–13.
[50] Samaritans and Swansea University, How Social Media Users Experience Self-Harm and Suicide Content, 2022; Samaritans, The Delve Study: Investigating User Journeys and Engagement with Online Self-Harm and Suicide Content, 2025, pp. 8–10.

[redacted]

platform-user groups to report encountering high-risk content on the relevant platform during the preceding week, at 4%; the same result was recorded among mid-level users who reported spending 30–120 minutes per day on the platform. The research nevertheless found that non-follower feeds were a significant reported source of exposure across all surveyed platforms and identified concerns regarding some Snapchat users' perceived ability to avoid repeated exposure to upsetting content. These findings are considered alongside the survey-based methodology and Snap's internal prevalence data.[51]

Likelihood conclusion

On the basis of the available external evidence, the platform-architecture factors recorded above, the moderation and recommendation controls applicable to this category, and the internal PVP, VVR and enforcement data, the likelihood of dissemination of content that glorifies self-harm on Snapchat's in-scope services is assessed at the lowest likelihood category. Under the methodology retained for year-on-year comparability, Self-Harm and Suicide-specific PVP increased from [redacted] recorded in the 2025 Report to [redacted] for the 90-day period ending 30 June 2026, while remaining very low in absolute terms. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP was [redacted] Considered together with the very low surface-specific VVR figures, including for EU minor users, and the other evidence set out above, these results support maintaining the **lowest likelihood classification** while warranting continued monitoring. There is no change in the determination from the 2025 Report.

Severity

Self-harm, suicide and eating disorders are recognised as among the most serious threats to individual physical and mental well-being. The World Health Organization has recorded suicide as a leading cause of death among young people globally; eating disorders are associated with substantial morbidity and elevated mortality risk; and self-injury, while distinct from suicidal behaviour, is associated with elevated subsequent suicide risk in the affected population. The exposure of vulnerable users, in particular minors, to content that glorifies, romanticises or instructs in respect of these behaviours can contribute to imitation effects, normalisation of harmful conduct, and barriers to help-seeking.

The severity classification for content glorifying self-harm, suicide or eating disorders is **serious harm industry-wide**, consistent with prior Reports. As a general matter, and consistent with Section 3.5, any specific instance presenting an imminent, credible threat to human life or to severe physical safety is treated as a Level 1 overall risk and handled accordingly through the

[51] Molly Rose Foundation, *Children's Exposure to Suicide, Self-Harm, and Eating Disorder Content on Social Media: Wave Two Findings*, June 2026, pp. 2, 6–8, 21, 61–62 and 70–73.

[redacted]

prioritised-response and emergency-escalation pathways recorded in the calibrated three-tier enforcement framework above.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend content that glorifies self-harm; no corresponding interest category exists. [redacted] Further detail in Section 5.6.
Content moderation systems	Proactive and [redacted] Reactive moderation is available through the in-app reporting flow under the Threats, Violence and Harm reporting category. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Threats, Violence and Harm prohibit the glorification of self-harm, including the promotion of self-injury, self-mutilation, suicide or eating disorders. [redacted] Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising content that glorifies, promotes or incites self-harm. Pre-publication review applies to all advertising. Further detail in Section 5.7.
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3, including in respect of generative-AI features the outputs of which could, in the absence of mitigations, produce content that glorifies self-harm.

[redacted]

Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of self-harm-related content on in-scope services is constrained by the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation). [redacted] Where glorifying content is identified (proactively or reactively), removal applies through the moderation pipeline described in Section 5.4.
Regional and linguistic aspects	The Here for You in-app resource portal is calibrated to surface region-specific mental-health resources and helpline routing for users searching for self-harm or suicide-related terms. Content is sourced from third-party mental-health organisations across multiple Member States.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, the default output for serious harm industry-wide combined with lowest likelihood is Level 3. However, applying the case-by-case deviation framework in Section 3.5, dissemination of content that glorifies self-harm is assessed at **Level 2** overall potential risk prioritisation. The upward adjustment reflects the potentially grave and lasting consequences of self-harm, suicide and eating-disorder content, including risks to life and physical and mental well-being, particularly for vulnerable users and minors. There is no change in this determination from the 2025 Report.

[redacted]

[redacted]

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. The platform architecture limits the structural opportunity for self-harm-glorifying content to reach a wide audience: recommended content in public feed is moderated; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation; no broad reshare functionality of Stories. The Here for You in-app resource portal surfaces mental-health resources when users search for self-harm-related terms or where Trust & Safety personnel determine intervention is appropriate. Resources are tailored to a user's geographic region. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Threats, Violence and Harm prohibit the glorification of self-harm, suicide or eating disorders. Enforcement is calibrated through the three-tier framework recorded in the Harm Description. Where content suggests an emergency situation involving imminent danger of death or serious bodily injury, response is prioritised, with proactive escalation to law enforcement or emergency authorities where appropriate. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Specific moderation procedures operate for self-harm-related content, [redacted] Help resources are included within rejection reasons surfaced to users when content is moderated. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend content glorifying self-harm; no corresponding interest category exists. [redacted] Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit advertising content that glorifies, promotes or incites self-harm. Further detail in Section 5.7.

[redacted]

Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for self-harm-related content is conducted on a regular cadence. The Self-Harm and Suicide-specific PVP result for the 90-day period ending 30 June 2026 and the six-month surface-specific VVR results for H1 2026 are recorded in the Likelihood section above, together with enforcement-volume and turnaround-time data. These results are monitored over time through the ongoing risk-detection processes described in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Limited application. Snap does not maintain a specific Article 22 DSA trusted-flagger relationship dedicated to self-harm and suicide content in the European Union; trusted-flagger relationships in the child-safety category may, in the ordinary course, address content overlapping with this risk.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap is a member of [redacted] an industry signal-sharing and best-practice initiative focused on suicide and self-harm content. Snap also engages with mental-health-focused civil-society organisations and academic experts on calibrated approaches to self-harm content moderation. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Threats, Violence and Harm category explainer (in the Community Guidelines Explainer Series), the Safety Center, and the Here for You in-app resource portal provide information on the substantive prohibitions, on enforcement, and on the resources available to users seeking support. The Here for You portal surfaces region-specific mental-health resources and helplines when users search for self-harm or suicide-related terms. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. [redacted]
Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines apply without regard to whether self-harm-related content is user-generated or AI-generated. Generative-AI feature outputs are subject to input/output safety scanning to prevent the generation of content that glorifies self-harm. Further detail in Section 5.9.

[redacted]

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of content that glorifies self-harm on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (recommended content in public feed is moderated, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation, recommendation-density adjustment for sensitive content), [redacted] the Here for You in-app resource portal with region-specific mental-health resources, the prioritised-response framework for emergency situations, the Thrive industry signal-sharing initiative, and the dedicated minor-protection framework support the conclusion that the mitigation framework operates effectively against this risk.

The mitigation framework will continue to be monitored on an ongoing basis through the Section 6 risk-detection processes, including in respect of the calibration of the enforcement framework, the operation of the Here for You portal, and the engagement with civil-society mental-health organisations and the Thrive initiative. There is no change in the overall conclusion from the 2025 Report.

4.1.9 Dissemination of Content encouraging or engaging in Violent or Dangerous Behaviour

This Section 4.1.9 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of content encouraging or engaging in violent or dangerous behaviour on Snapchat's in-scope services.

Harm Description

This category covers violence (in its physical and psychological dimensions, beyond the harassment-and-bullying context covered in Section 4.1.7) and risky or dangerous behaviours and activities. The category encompasses, without limitation:

- **Violent, threatening or abusive behaviour.** Content expressing an intent to inflict serious physical or emotional harm on an individual or group, or damage to their property, beyond the harassment-and-bullying scope of Section 4.1.7.
- **Threats and intimidation.** Explicit or implied threats - of violence, retaliation, or psychological harm - including content that attempts to coerce or manipulate through fear, blackmail or control.

[redacted]

- **Coercive and controlling behaviour.** Content that reflects or promotes coercive tactics, including manipulation, isolation, surveillance or restriction of a person's freedom or autonomy.
- **Vigilante activity.** Coordinated efforts to intimidate or take physical action against individuals or communities outside of proper legal process.
- **Encouraging or engaging in dangerous behaviour.** Content that promotes or participates in risky activities likely to be imitated and that could result in serious harm - including reckless driving, unsafe "challenges" and behaviour that endangers personal or public safety.
- **Glorification or incitement of harmful behaviour.** Content that glorifies or encourages violence or abuse, including towards persons or animals, including Snaps depicting gratuitous violence, graphic injury, animal cruelty or other distressing content.

The substantive prohibitions in respect of this category are set out in the Community Guidelines on Threats, Violence and Harm, the corresponding Community Guidelines Explainer, and the Snap Transparency Report Glossary. Certain activities in this category - including the sharing of credible, imminent threats to human life or to severe physical safety - are designated severe harm in the Community Guidelines and are subject to swift enforcement consequences.

Likelihood

Internal data

Prevalence Testing for this category is conducted across two sub-categories: Violent and Disturbing Content, covering the violent, glorification and graphic-injury dimensions; and Dangerous Activities, covering risky or potentially imitative behaviour. Prevalence figures on Snapchat's in-scope services[52] for the 90-day period ending 30 June 2026 are recorded below:



[52] Spotlight, Public Story, Maps, Saved Story, Communities.

[redacted]



The 2024 Report identified a Q3 2023 increase in Violent and Disturbing Content prevalence, in response to which a new proactive-detection mechanism was deployed. Following deployment, prevalence declined materially, from approximately [redacted] before deployment to [redacted] in July 2024 and [redacted] in April 2025. Using the methodology retained for year-on-year comparability in Section 6.4, the 2026 Violent and Disturbing Content PVP result was [redacted] for the 90-day period ending 30 June 2026, compared with [redacted] recorded in the 2025 Report. This represents a slight increase under the comparable initial-review methodology, while prevalence remained substantially below the [redacted] recorded in July 2024 and the approximately [redacted] level observed before deployment of the proactive-detection mechanism. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP was [redacted] The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction. The continued operation of the proactive-detection mechanism therefore remains an important component of Snap's content-moderation framework.

Transparency Report enforcement data

The H2 2025 global Transparency Report recorded 1,088,184 content and account reports under the Threats and Violence policy category. Snap recorded 153,715 enforcements involving 127,252 unique accounts, with a median turnaround time of one minute from detection to final action. Although these figures are global and platform-wide, rather than limited to EU users or Snapchat's in-scope services, they provide additional context on the operational scale and speed of enforcement for this category.

Violative View Rate (H1 2026)

Six-month average VVR data for the relevant enforcement categories for EU users during the period from 1 January to 30 June 2026 are set out below for Spotlight, Public Story and Maps:

[redacted]

[redacted]

The corresponding six-month average VVR for EU minor users aged 13–17 was as follows:

[redacted]

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. For Lenses, during H1 2026, submissions were rejected under broader operational categories relevant to this risk, including promotion of dangerous, harmful or illegal activity [redacted] and content inciting or glorifying violence or the use of weapons [redacted]. These figures are Lens-review rejection volumes and should not be interpreted as PVP or VVR measures.

The results indicate that violative views remained low in absolute terms across the assessed in-scope services. The minor-user results were lower than the corresponding all-user results across each assessed sub-category and surface. Violence recorded the highest VVR within this category on Spotlight, while Dangerous Activities recorded the highest VVR on Public Stories.

External evidence

Recent peer-reviewed research continues to identify associations between aspects of adolescent social-media use and aggressive behaviour, while emphasising that the relationship is shaped by multiple individual, psychological and contextual factors and does not support a simple causal

[redacted]

pathway. A 2025 scoping review of 44 quantitative studies involving more than 90,000 adolescents across 19 countries identified a range of mediating and moderating factors, including cyberbullying involvement, problematic use, peer and family context, and individual psychological characteristics, relevant to the relationship between social-media use and aggression.[53]

Separately, a recent peer-reviewed scoping review of risky social-media challenges identified a limited but growing evidence base documenting injuries associated with such challenges, including burns, poisoning, lung injury and strangulation. The review noted that the available literature was predominantly observational and that further research is required regarding participation, causation and the effectiveness of interventions.[54]

Likelihood conclusion

On the basis of the internal Prevalence Testing data, the low absolute prevalence of both PVP sub-categories, and the low VVR figures across the assessed in-scope services, including for EU minor users, the likelihood of dissemination of violent or dangerous content on Snapchat's in-scope services is assessed at the lowest likelihood category. Under the methodology retained for year-on-year comparability, Violent and Disturbing Content increased slightly from [redacted] in the 2025 Report to [redacted] for the 90-day period ending 30 June 2026, while Dangerous Activities decreased from [redacted] to [redacted]. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP figures were [redacted] and [redacted] respectively. Considered together with the longer-term prevalence trend, the VVR results and the other evidence set out above, these results do not alter the existing lowest-likelihood determination. There is no change in this determination from the 2025 Report.

Severity

The severity of harm associated with this category varies materially by the specific instance:

- Content involving imminent credible threats to human life or to severe physical safety, including, for example, school-shooting threats, bomb threats, missing-person cases, or specific direct threats of violence, represents severe harm. Such content is treated as a Level 1 overall risk under the Section 3.5 case-by-case framework and triggers prioritised response, including (where appropriate) proactive escalation to law enforcement. School-shooting and similar incidents have, regrettably, demonstrated the potential for severe and lasting consequences for victims, witnesses and broader communities;
- Viral "challenges" and dangerous-activity content can vary in severity from limited (comedic pratfalls, minor injury depictions) to serious (challenges that have resulted in

[53] Georgios Giannakopoulos and Afroditi Prassou, Mediating and Moderating Mechanisms in the Relationship Between Social Media Use and Adolescent Aggression: A Scoping Review of Quantitative Evidence, European Journal of Investigation in Health, Psychology and Education, Vol. 15, Issue 6, 2025, Article 98.
[54] Leah K. Middelberg et al., Risky Social Media Challenges: A Scoping Review, 2000–2024, Injury Epidemiology, Vol. 13, Article 1, 2026.

[redacted]

injury or death). Where a specific challenge or dangerous-activity pattern emerges with the potential to spread and cause material harm, the Content Crisis Response Protocol may be activated;

- Threats, intimidation, and coercive control content represent harms that vary by intensity and target context, and that overlap with the harassment-and-bullying scope of Section 4.1.7 and (in some cases) the sexual-crimes scope of Section 4.1.6;
- Glorification or incitement content (including gratuitous violence and animal-cruelty content) represents harm to public morals and can have specific impact on minor users encountering such content.

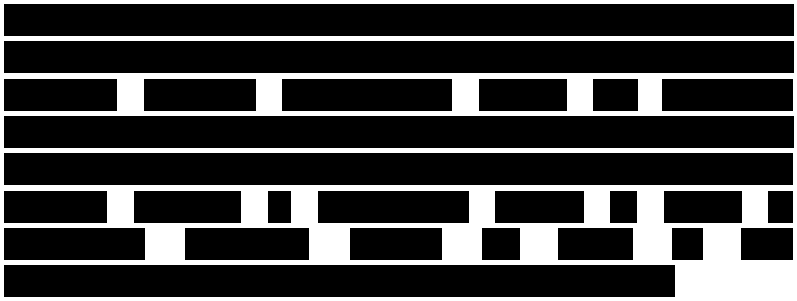
In recognition of the variable severity profile across these sub-categories, and the moderating effect of the operational mitigation framework on the most severe sub-categories, the overall severity classification for this category is **significant harm industry-wide**, consistent with prior Reports. As a general matter, and consistent with Section 3.5, any specific instance presenting an imminent, credible threat to life or severe physical safety is treated as a Level 1 overall risk and handled accordingly.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend content encouraging or engaging in violent or dangerous behaviour; no corresponding interest category exists. The absence of an open recommender feed for unvetted user-generated content, the absence of a broad reshare functionality of Stories, and pre-distribution moderation for content reaching wide audiences reduce the structural opportunity for amplification of violent or dangerous content (including viral "challenges"). Further detail in Section 5.6.
Content moderation systems	[redacted]

[redacted]

Applicable terms and conditions and enforcement	The Community Guidelines on Threats, Violence and Harm prohibit the categories described in the Harm Description above. Certain activities falling within this category, including the sharing of credible, imminent threats to human life, safety or well-being, and specific threats of violence or other serious criminal activities, are categorised as severe harm in the Community Guidelines and trigger swift enforcement consequences as described in the Severe Harms explainer. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising containing violent, threatening or dangerous content. Pre-publication review applies to all advertising. Further detail in Section 5.7.
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3.
Intentional manipulation (Article 34(2))	
Amplification and rapid, wide dissemination	Amplification of violent or dangerous content on in-scope services is constrained by the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation for content reaching wide audiences). Where violent or dangerous content is identified proactively or reactively, enforcement action is taken rapidly. The H2 2025 global Transparency Report recorded a median turnaround time of one minute from detection to final action for the Threats and Violence policy category. Although this figure is global and platform-wide rather than limited to EU users or Snapchat's in-scope services, it provides relevant context on the speed of enforcement for this category.

[redacted]

Regional and linguistic aspects	ALD banned-word lists for violent and threat-related content are maintained across multiple languages and categorised by context and region. Moderation guidance is calibrated to Member State legal contexts where required (for example, in respect of categories of cyber-violence, stalking and intimidation that are subject to specific Member State criminal legislation). The Content Crisis Response Protocol applies to elevated-risk events with potential Member State-specific dimensions.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, dissemination of content encouraging or engaging in violent or dangerous behaviour is assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of significant harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to limit the structural opportunity for violent or dangerous content distribution: recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences. The Content Guidelines for Recommendation Eligibility limit further the categories of content eligible for algorithmic distribution, beyond the substantive Community Guidelines prohibitions. Further detail in Section 5.1.

[redacted]

Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Threats, Violence and Harm prohibit the categories described in the Harm Description. Where content suggests an emergency situation involving an imminent, credible threat to human life or to severe physical safety, response is prioritised, with proactive escalation to law enforcement available where appropriate [redacted] Where content is identified as violating during enforcement, content is, in the ordinary course, retained for an extended period to support law-enforcement investigation. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. [redacted]
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend violent or dangerous content; no corresponding interest category exists. The Content Guidelines for Recommendation Eligibility set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit violent, threatening or dangerous content in advertisements. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing is conducted on a regular cadence for both Violent and Disturbing Content and Dangerous Activities sub-categories; results are recorded in this Report and in Snap's Transparency Reports. Where prevalence variations are observed (including the early-2025 Dangerous Activities variation referenced in the Likelihood discussion above), the underlying cause is investigated and corrective action taken where appropriate. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies, in limited form. Notices submitted by designated trusted flaggers under Article 22 DSA may concern violent or dangerous content and are processed through the applicable priority-notice procedures. Orders issued by competent authorities under Articles 9 and 10 DSA are processed through the corresponding DSA order-handling pathway. Snap also cooperates separately with Member State internet referral units and law-enforcement bodies in

[redacted]

	respect of imminent credible threats through the Law Enforcement Operations team. Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap operates an internal Content Crisis Response Protocol (Section 5.12) that applies in the event of major incidents [redacted] Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Threats, Violence and Harm category explainer (in the Community Guidelines Explainer Series), the Severe Harms explainer, the Safety Center, and Snap's Transparency Reports provide information on the substantive prohibitions, on enforcement and prevalence, and on the resources available to users. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. [redacted]
Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines apply without regard to whether violent or dangerous content is user-generated or AI-generated, including in respect of synthetic content depicting violence. Generative-AI Lens flows include input/output scanning, retry logic for unsafe outputs, and watermarking of generative-AI outputs (Section 5.9).

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of content encouraging or engaging in violent or dangerous behaviour on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation), the proactive moderation suite [redacted], the swift enforcement of severe-harm sub-categories with

law-enforcement escalation pathways, the Content Crisis Response Protocol for elevated-risk events, and the Family Center oversight for minor users supports the conclusion that the mitigation framework operates effectively against this risk. The low absolute prevalence of Violent and Disturbing Content, the substantial reduction in Dangerous Activities PVP, and the low VVR figures, including for EU minor users, support this conclusion.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to the emergence of any new viral-challenge formats or sub-category patterns that may warrant a targeted operational response.

4.1.10 Dissemination of Harmful False Information

This Section 4.1.10 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of Harmful False Information on Snapchat's in-scope services. This risk is closely related to, and should be read alongside, Section 4.3.1 (Negative Effects on Democratic and Electoral Processes) and Section 4.3.2 (Negative Effects on Civic Discourse), which engage the broader systemic risk to civic and democratic functioning of which Harmful False Information is a contributing factor.

Harm Description

Snap describes Harmful False Information as content that is false or misleading and that causes harm or is malicious. The category covers the following sub-types of false-information content, recognising that these sub-types overlap and that the operational mitigation framework applies uniformly across them:

- **Disinformation.** Content that is deliberately false and designed to cause harm.
- **Misinformation.** Content that is false or misleading in ways that can cause harm, shared without knowledge that the information is false.
- **Mal-information.** Genuine information shared with intent to cause harm (for example, the leaking of private communications with malicious intent).

Examples of conduct that, in the ordinary course, fall within this category include:

- sharing false claims about voting dates, ballot access or eligibility requirements;
- impersonating a public figure with the intent to deceive;
- denying major tragic events (for example, the Holocaust or specific identified mass-violence incidents);
- disseminating unsubstantiated medical claims (including in connection with vaccine misinformation, untested therapies, or so-called "conversion therapy");
- deepfakes or AI-manipulated content with misleading implications, including content depicting real persons or events that do not represent authentic source material.

[redacted]

The category also extends to coordinated or sponsored activity by domestic or foreign-state actors that aims to manipulate the information environment, including in connection with elections, public-health events, or other public-interest moments. The substantive prohibitions are recorded in the Community Guidelines on Harmful False or Deceptive Information, the corresponding Community Guidelines Explainer, and the Snap Transparency Report Glossary.

Likelihood

Internal data

Prevalence Testing for Harmful False Information continued to record extremely low prevalence on Snapchat's in-scope services[55] for the 90-day period ending 30 June 2026. Harmful False Information PVP was recorded as [redacted] in H1 2026 at the displayed precision, compared with [redacted] recorded in the 2025 Report.

Six-month average VVR data for Harmful False Information for EU users during the period from 1 January to 30 June 2026 are set out below:

[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. For Lenses, during H1 2026, 2 submissions were rejected under the operational category "Misinformation/Disinformation." This figure is a Lens-review rejection volume and should not be interpreted as a PVP or VVR measure.

Snap's H2 2025 global Transparency Report recorded 5,202 enforcement actions against False Information content, affecting 5,081 unique accounts, with a median turnaround time of one minute from detection to final enforcement action.

Election-cycle evidence

The 2024 European Parliament elections provided a substantive operational test of Snap's mitigation framework in respect of Harmful False Information bearing on civic and electoral processes:

- Snap participated in the European Commission's stress test and multi-stakeholder roundtable dialogues ahead of the elections.

[55] Spotlight, Public Story, Maps, Saved Story, Communities.

- As reported in Snap's election-integrity update of 24 June 2024, Snap observed a small increase in reported activity during the election period but did not identify any material incidents or threats.
- Snap's moderation and reporting tools operated as designed, and none of the content reported during the election period was verified as misinformation on Snapchat.
- Snap previously participated in the Election Integrity Partnership threat-escalation programme in connection with the 2022 election cycle. During that period, Snap received one incident report from EIP researchers monitoring risks on Snapchat.

External evidence

External evidence continues to identify Harmful False Information as a material industry-wide risk. A 2024 European Parliamentary Research Service briefing identifies information manipulation as affecting public opinion, including in electoral contexts, and records its increasingly visible impacts on societies, democracy, health, security and human rights. It further identifies the potential for information manipulation to undermine democratic resilience and benefit actors seeking to weaken trust in democratic systems.[56]

More recent European Parliament research has identified generative-AI tools as an emerging information-integrity risk, including by making the production and dissemination of deceptive text, images, audio and video easier, cheaper and more scalable.[57]

Likelihood conclusion

On the basis of the internal Prevalence Testing data, the very low surface-specific VVR figures for EU users and EU minor users, the operational performance during the 2024 European Parliament elections, the rapid median turnaround time for enforcement action, and the platform-architecture factors recorded below, the likelihood of dissemination of Harmful False Information on Snapchat's in-scope services is assessed at the lowest likelihood category. Harmful False Information PVP was recorded as [redacted] in H1 2026 at the displayed precision, while Snap's H2 2025 global Transparency Report recorded a median turnaround time of one minute from detection to final enforcement action. These results support maintaining the **lowest likelihood classification**. There is no change in this determination from the 2025 Report.

Severity

Harmful False Information does not, in the ordinary course, present an imminent threat to human life. The category is, however, recognised as posing material risk to public discourse, democratic

[56] European Parliamentary Research Service, Naja Bentzen, Online Information Manipulation and Information Integrity: An Overview of Key Challenges, Actors and the EU's Evolving Response, 30 September 2024, PE 762.416.
[57] European Parliamentary Research Service, Naja Bentzen, Information Manipulation in the Age of Generative Artificial Intelligence, December 2025, PE 779.259.

functioning, public health and individual well-being. The severity of harm associated with this category varies considerably across sub-types:

- Election-related and civic-process disinformation can undermine the integrity of electoral processes and the legitimacy of democratic institutions. The associated harms are addressed in detail in Section 4.3.1;
- Public-health misinformation (including in connection with vaccines, untested therapies, and emergency public-health events) can lead to direct harm to individuals through encouragement of unsafe behaviour or refusal of effective interventions. The associated harms are addressed in detail in Section 4.4.1 (Public Health);
- Tragic-event denial and historical-fact denial cause psychological harm to affected communities and contribute to the broader normalisation of harmful ideologies; this overlaps with Section 4.1.2 (Hate Speech) where the denial constitutes hateful conduct;
- Impersonation-based misinformation can cause direct harm to the impersonated individual (including reputation harm and, in some cases, financial harm to those deceived);
- Synthetic content (deepfakes, AI-manipulated content) presents an emerging risk dimension that overlaps with the content-authenticity measures recorded in Section 5.9.

Recognising both the variable severity profile and the moderating effect of the platform-architecture and operational mitigation framework, the overall severity classification for this category is **significant harm industry-wide**, consistent with prior Reports. As a general matter, and consistent with Section 3.5, any specific instance presenting an imminent, credible threat (for example, in connection with a public-health emergency or an active election integrity threat) is treated as a Level 1 overall risk and handled accordingly.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend Harmful False Information; no corresponding interest category exists. Steps are taken to limit the surfacing of misleading or sensationalist headlines. The absence of an open recommender feed for unvetted user-generated content, the absence of a broad reshare functionality of Stories, and pre-distribution moderation for content reaching wide audiences reduce the structural opportunity for amplification of false or misleading content. Further detail in Section 5.6.

[redacted]

Content moderation systems	Proactive and reactive moderation procedures operate to detect and remove Harmful False Information. [redacted] Snapchatters can report Harmful False Information through the in-app reporting flow under the dedicated False Information reporting reason. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Harmful False or Deceptive Information prohibit content that is false or misleading and causes harm or is malicious, including the categories enumerated in the Harm Description above. The Community Guidelines apply equally to all users, including politicians, public figures and Media Partners. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Every political, health-related and sensitive-issue advertisement is reviewed by a human ad-review team. Advertisements containing unsubstantiated or false claims are rejected. Political advertising in the European Union has been discontinued as of October 2025, consistent with the EU's Political Advertising Regulation (further detail in Section 4.3.1 and Section 5.7); the political advertising library established under prior practice continues to be made available for historical disclosure. Further detail in Section 5.7.
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3, with particular attention to generative-AI features the outputs of which could surface synthetic misinformation (deepfakes, AI-manipulated images, voice cloning).
Intentional manipulation (Article 34(2))	[redacted]

[redacted]

	[redacted]
Amplification and rapid, wide dissemination	Amplification of Harmful False Information on in-scope services is constrained by the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation). Where Harmful False Information is identified (proactively or reactively), the median turnaround time for enforcement action is rapid; Snap's H2 2025 global Transparency Report recorded a median turnaround time of one minute from detection to final enforcement action.
Regional and linguistic aspects	ALD banned-word lists and misinformation classifiers are maintained across multiple languages. Member State election cycles are monitored individually by the Trust & Safety team and the Content Crisis Response Protocol may be activated for elections with elevated risk profiles. Engagement with Member State regulators (notably the French regulator Arcom, which monitors industry action against misinformation) is maintained on a continuing basis.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, dissemination of Harmful False Information is assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of significant harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.



Snap's Mitigations

[redacted]

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to limit the structural opportunity for amplification of false or misleading content: recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences. Snap's product design and policy practices substantially limit virality, which itself reduces the impact of false content. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Harmful False or Deceptive Information prohibit the categories enumerated in the Harm Description. The Community Guidelines apply equally to all users, including public figures, politicians, athletes and Media Partners; no exception is granted on the basis of newsworthiness or public interest. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. [redacted] Reactive moderation is available through the in-app reporting flow with a dedicated False Information reporting reason. Snap's H2 2025 global Transparency Report recorded a median turnaround time of one minute from detection to final enforcement action for this category. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend Harmful False Information; no corresponding interest category exists. Steps are taken to prevent surfacing of misleading or sensationalist headlines. The Content Guidelines for Recommendation Eligibility set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. The Advertising Policies prohibit advertising containing unsubstantiated or false claims. Every political, health-related and sensitive-issue advertisement is reviewed by a human ad-review team, with advertisements containing unsubstantiated or false claims rejected. Political advertising in the European Union has been discontinued as of October 2025, consistent with the EU's

[redacted]

	Political Advertising Regulation; this discontinuation is recorded in detail in Section 4.3.1 and Section 5.7. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Harmful False Information is conducted on a regular cadence. Harmful False Information PVP was recorded as [redacted] in H1 2026 at the displayed precision, and the corresponding surface-specific VVR figures for EU users and EU minor users are recorded in the Likelihood section above. Snap's Transparency Reports separately provide enforcement-volume and turnaround-time information for False Information and Impersonation. Further detail is provided in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies, in limited form. Trusted flaggers active in other Community Guidelines categories may also report Harmful False Information; in practice, this is uncommon given the limited prevalence of such content on the in-scope services.
Codes and Crisis Protocols - Art. 35(1)(h)	Limited application. As at the 2025 Report, Snap is not a signatory of the EU Code of Practice on Disinformation (now operating as a Code of Conduct under the DSA framework as of 1 July 2025). The 2025 Report recorded that Snap had limited exposure to the underlying risk, and that Snap had directed resources to other codes more closely aligned with the in-scope risk profile. Snap engages with [redacted] on a continuing basis in respect of monitoring industry action against misinformation. Snap also participated in the European Commission's stress test and multi-stakeholder roundtable dialogues ahead of the 2024 European Parliament elections, and operates an internal Content Crisis Response Protocol (Section 5.12).
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Harmful False or Deceptive Information category explainer (in the Community Guidelines Explainer Series), the Safety Center, and Snap's Transparency Reports provide information on the substantive prohibitions, on enforcement and prevalence, and on the resources available to users. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. [redacted]

[redacted]

	resources made available through the Safety Center. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. The Community Guidelines apply without regard to whether Harmful False Information is user-generated or AI-generated, including in respect of synthetic misinformation (deepfakes, AI-manipulated content). The content-authenticity measures described in Section 5.9 (watermarking, contextual identification of generative-AI content) operate alongside the substantive prohibitions. Generative-AI features made available on Snapchat operate within content-policy and rights-handling frameworks designed to mitigate the risk that their outputs could surface deceptive synthetic content. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of Harmful False Information on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation), the proactive moderation suite [redacted] the rapid median turnaround time for enforcement action of one minute, the political-advertising discontinuation in the European Union, the engagement with Member State regulators (notably Arcom), and the Content Crisis Response Protocol for elevated-risk events supports the conclusion that the mitigation framework operates effectively against this risk. The successful operational performance during the 2024 European Parliament elections supports this conclusion.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to the emergence of synthetic-content (deepfake and AI-manipulated) misinformation patterns and to events presenting an elevated risk profile (including elections in Member States and other public-interest events).

4.1.11 Dissemination of Fraud and Spam

This Section 4.1.11 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of Fraud and Spam content on Snapchat's in-scope services.

[redacted]

Harm Description

Fraud and Spam, for the purposes of this assessment, encompass two related but distinguishable sub-categories:

- **Fraud.** Content that constitutes scams or deceptive marketing practices that abuse community trust or entice users to make purchases or engage in activities under misleading pretences. Fraud sub-types include financial scams, identity misrepresentation, fake investment offers, romance scams, phishing, and other deceptive practices.
- **Spam.** Content involving fake or artificially-inflated engagement, including pay-for-follower schemes and other follower-growth schemes; the creation or sharing of content or services that mimic the appearance or function of Snapchat features, formats or advertisements in order to gain clicks or follows; the promotion of spam applications; or mass untargeted, repetitive or frequent posting or sharing.

The substantive prohibitions are recorded in the Community Guidelines on Harmful, False or Deceptive Practices, the Advertising Policies, the corresponding Community Guidelines Explainer, and the Snap Transparency Report Glossary. Fraud-related conduct can also overlap with categories addressed elsewhere in this Report, including the sextortion sub-category in Section 4.1.6 (where financial sextortion combines fraud with sexual-crime elements) and the impersonation patterns in Section 4.1.10 (where impersonation is used to disseminate Harmful False Information).

Likelihood

Internal data - content side

Prevalence Testing for Fraud and Spam content recorded a PVP of [redacted] on Snapchat's in-scope services[58] for the 90-day period ending 30 June 2026, compared with [redacted] recorded in the 2025 Report. The 2026 result therefore represents a further decrease relative to the prior assessment cycle and remains substantially below the [redacted] recorded in August 2023.

[redacted]

[58] Spotlight, Public Story, Maps, Saved Story, Communities.

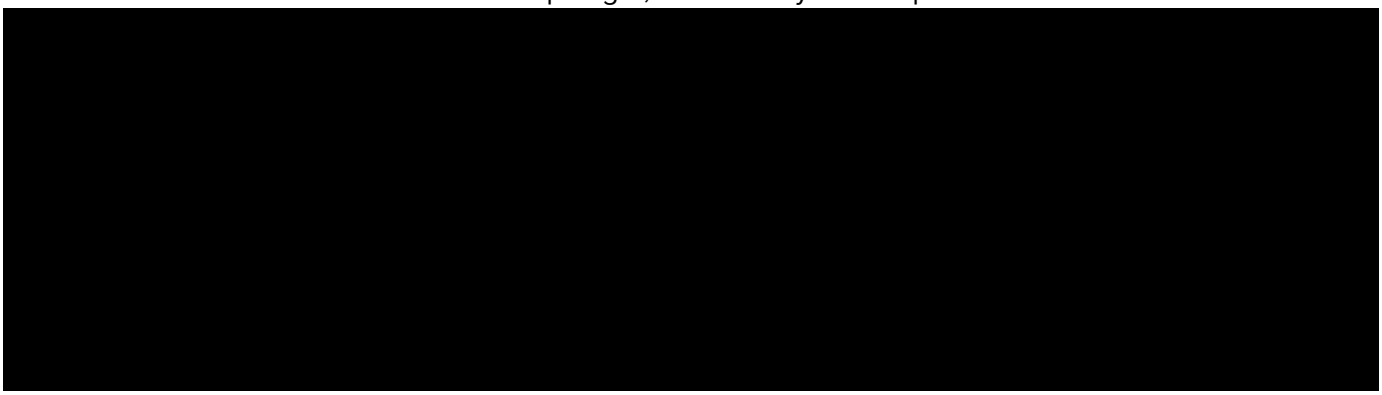
[redacted]

Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP for the 90-day period ending 30 June 2026 was [redacted] The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction.

The 2025 Report recorded the movement of Fraud and Spam from medium likelihood to lowest likelihood. The further decrease in the 2026 PVP snapshot is considered below alongside the surface-specific VVR results, the advertising-review data, the applicable platform-design factors and the broader mitigation framework in confirming whether the lowest likelihood classification remains appropriate.

Violative View Rate (H1 2026)

Six-month average VVR data for Fraud and Spam for EU users during the period from 1 January to 30 June 2026 are set out below for Spotlight, Public Story and Maps:



The corresponding six-month average VVR data for EU minor users aged 13–17 were as follows:



Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets rather than VVR. Advertising-side data are set out below. The Lens review taxonomy does not contain a Fraud and Spam-specific rejection category, so no directly comparable Lens rejection volume is presented in this section.

Internal data - advertising side

[redacted]

[redacted] The volumes for the most recent reporting period and the comparable prior period are recorded below:[59]

[redacted]

Fraud- and schemes-related rejection reasons continued to account for material volumes of advertising rejections during the assessment period. Between 1 July 2025 and 30 June 2026, [redacted] advertisements were rejected, compared with [redacted] during the preceding 12-month period, representing a decrease of approximately [redacted] Over the same period, the number of distinct advertisements completing final review increased from [redacted] an increase of approximately [redacted] These figures demonstrate the expanded operational scale of the advertising-review pipeline.

For additional context, the post-publication advertising reporting volumes recorded for H2 2025 were as follows:

Post-publication ad reporting (1 July-31 December 2025)

[redacted]

The H2 2025 Transparency Report recorded 81,976 advertisements reported following publication, of which 47,266 were removed. These figures cover all advertising-policy reasons and provide additional context regarding the scale of Snap's post-publication reporting and enforcement processes alongside the substantially larger pre-publication review pipeline.

External evidence

External evidence relevant to this risk includes:

- **UK banking-sector evidence.** Evidence provided by TSB to the UK Home Affairs Committee in February 2024 indicated that approximately 80% of the principal authorised

[59] Snap Inc., internal advertising-review pipeline data covering 1 July 2024 to 30 June 2026, including total advertisements completing final review, total advertisements rejected, and prior-period comparators.

push-payment fraud types affecting its customers began on social-media services, with the majority originating on Meta-owned platforms. More recent TSB customer data published in 2025 concerning ticket scams attributed 66% of cases to Facebook, 19% to X, 5% to Instagram, 4% to Snapchat and 4% to TikTok. Although the later dataset is limited to ticket scams and should not be treated as a platform-wide measure of fraud prevalence, it provides more recent directional evidence that Snapchat accounted for a comparatively small share of the relevant TSB cases[60].

- **ESMA and IOSCO action on online financial misconduct**. On 28 May 2025, the European Securities and Markets Authority wrote to Snap and several other online-platform providers encouraging them to take proactive steps to prevent the promotion of unauthorised financial services. ESMA identified online scams targeting retail investors as a serious risk capable of causing financial loss and undermining trust in the financial sector. The action followed IOSCO's 21 May 2025 Statement on Combatting Online Harm and the Role of Platform Providers, which addressed the role of online platforms in limiting financial misconduct. These materials establish the increasing regulatory focus on platform measures addressing online financial scams; they do not, without separate internal confirmation, establish that Snap introduced any particular operational measure in response[61].
- **Alan Turing Institute research on AI-enabled crime.** A March 2025 report by the Center for Emerging Technology and Security at the Alan Turing Institute found that, although criminal use of AI remained at an early stage, evidence indicated a substantial acceleration in AI-enabled crime, particularly in financial crime, phishing and romance scams. The report identified the ability of AI systems to automate and rapidly scale existing criminal activity and to exploit victims' psychological vulnerabilities. The findings provide relevant context for the emerging AI-enabled fraud patterns considered in this assessment but do not provide a Snapchat-specific measure of prevalence.[62]
- **Europol cybercrime threat assessment.** Europol's 2026 Internet Organised Crime Threat Assessment identified automation and artificial intelligence as increasingly important enablers of online fraud. Europol reported that generative-AI tools are being used to tailor social-engineering techniques and increase the efficiency and scale of fraud schemes, alongside facilitators such as caller-ID spoofing and SIM farms capable of distributing large volumes of messages, calls and social-media posts. These findings support

[60] House of Commons Home Affairs Committee, Oral Evidence: Fraud, HC 125, 7 February 2024, Q151–152; TSB, "Oasis and Taylor Swift fans lose out to scammers, TSB data reveals" (2025).
[61] European Securities and Markets Authority, "ESMA urges social media companies to tackle unauthorised financial ads" (28 May 2025), including Letter ESMA35-1872330276-2397 to Snap; International Organization of Securities Commissions, "Statement on Combatting Online Harm and the Role of Platform Providers" (21 May 2025).
[62] Joe Burton, Ardi Janjeva, Simon Moseley and Alice, AI and Serious Online Crime, Center for Emerging Technology and Security, Alan Turing Institute (31 March 2025).

[redacted]

continued monitoring of evolving AI-enabled and large-scale fraud methods, but do not provide a Snapchat-specific measure of prevalence.[63]

Likelihood conclusion

On the basis of the internal data, including the PVP result of [redacted] for the 90-day period ending 30 June 2026 under the methodology retained for year-on-year comparability, the low surface-specific VVR results across the assessed in-scope services, and the advertising-review data, the likelihood of dissemination of Fraud and Spam on Snapchat's in-scope services is assessed at the lowest likelihood category. PVP decreased from the [redacted] result recorded in the 2025 Report and remains substantially below the [redacted] recorded in August 2023. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP was [redacted] Between 1 July 2025 and 30 June 2026, [redacted] distinct advertisements completed final review and [redacted] advertisements were rejected. The H2 2025 Transparency Report separately recorded 81,976 advertisements reported following publication and [redacted] removed across all advertising-policy reasons. Considered alongside the low VVR results, the platform-design constraints, the applicable mitigation framework and the external evidence, including recent TSB data indicating that Snapchat accounted for a comparatively small share of the purchase-fraud cases in that dataset, the PVP results further support the lowest-likelihood determination. There is no change in this determination from the 2025 Report.

Severity

Fraud and Spam do not, in the ordinary course, present an imminent, credible threat to human life or severe physical safety. They can nevertheless cause direct financial loss, unauthorised access to accounts, privacy and cybersecurity harms, reputational and identity harm, and broader harm to users' confidence in legitimate online activity.

The severity of an individual incident may vary depending on the scale of the loss, the vulnerability of the affected user and the nature of the underlying conduct. Where fraud overlaps with sexual exploitation, including sextortion, or specifically targets minors, the elevated severity of that conduct is addressed in the relevant risk descriptions, including Sections 4.1.6 and 4.4.3. Any specific instance presenting an imminent, credible threat to life or serious physical safety is addressed through the case-by-case framework described in Section 3.5.

Taking these factors into account, the severity of the negative effects associated with Fraud and Spam is assessed as significant harm industry-wide, consistent with prior Reports.

DSA Risk Factors

[63] Europol, The Evolving Threat Landscape: How Encryption, Proxies and AI Are Expanding Cybercrime - Internet Organised Crime Threat Assessment 2026 (28 April 2026).

[redacted]

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend Fraud and Spam content; no corresponding interest categories exist. Steps are taken to limit the surfacing of artificially-inflated engagement signals (including pay-for-follower content, follower-growth schemes and content mimicking Snapchat features). The absence of an open recommender feed for unvetted user-generated content reduces the structural opportunity for amplification of fraud or spam content. Further detail in Section 5.6.
Content moderation systems	Proactive and reactive moderation procedures operate to detect and remove Fraud and Spam content. [redacted] Snapchatters can report Fraud and Spam content through the in-app reporting flow under the dedicated reporting reasons. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Harmful, False or Deceptive Practices prohibit fraud, scams, and deceptive activity, including false claims, impersonation with intent to deceive, and engagement manipulation. The Advertising Policies prohibit fraud-related advertising practices, including unsubstantiated claims, deceptive offers and misleading business practices. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	All advertisements are subject to pre-publication review against the Advertising Policies, which prohibit fraud-related categories. Higher-risk advertising categories [redacted] are subject to additional human review. [redacted] further detail in the Likelihood discussion above and in Section 5.7.
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3, including in respect of generative-AI features the outputs of which could be used to facilitate fraud (for example, AI-generated impersonation content, deepfake voice cloning for vishing scams).

[redacted]

Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of Fraud and Spam content on in-scope services is constrained by the platform architecture, including the absence of an open recommender feed for unvetted user-generated content and broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), and pre-distribution moderation, and by the pre-publication review pipeline that applies to all advertising. Between 1 July 2025 and 30 June 2026, [redacted] distinct advertisements completed final review and [redacted] advertisements were rejected across all advertising-policy reasons. These figures demonstrate the operational scale of the review pipeline but do not support attributing all rejected advertisements specifically to Fraud and Spam. Further detail is provided in the Likelihood analysis above.
Regional and linguistic aspects	ALD banned-word lists for fraud and spam content are maintained across multiple languages. Cross-industry and law-enforcement engagement on cross-border fraud patterns is maintained, including in connection with Member State and EU-level financial-supervision frameworks.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with the 2025 Report, dissemination of Fraud and Spam is assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of **significant harm industry-wide** with **lowest likelihood**, without case-by-case deviation. The 2025 Report recorded the movement from medium likelihood to lowest likelihood. There is no further change in the 2026 determination.

[redacted]

[redacted]

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to limit the structural opportunity for fraud and spam distribution: recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences. The minor-account default settings (Section 5.1.3) reduce minor exposure to fraud-related contact pathways. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Harmful, False or Deceptive Practices prohibit fraud, scams, deceptive activity and engagement manipulation. The Advertising Policies prohibit fraud-related advertising practices. Account-level enforcement applies for repeat or coordinated fraud activity. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. [redacted] Reactive moderation is available through the in-app reporting flow with dedicated reporting reasons. [redacted] Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend Fraud and Spam content; no corresponding interest categories exist. Steps are taken to limit surfacing of artificially-inflated engagement signals. Further detail in Section 5.6.

[redacted]

Advertising Systems - Art. 35(1)(e)	Applies. The Advertising Policies prohibit fraud-related advertising. [redacted] Between 1 July 2025 and 30 June 2026, [redacted] distinct advertisements completed final review and [redacted] advertisements were rejected across all advertising-policy reasons. The number of advertisements completing final review increased by approximately [redacted] compared with the preceding 12-month period, while total rejection volume decreased by approximately [redacted] Post-publication reporting and enforcement pathways also apply, with the relevant H2 2025 volumes recorded in the Likelihood analysis above. Further detail is provided in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Fraud and Spam is conducted on a regular cadence; ad review and rejection volumes, post-publication report volumes, and content-side enforcement metrics are all monitored on a continuing basis. Results are recorded in this Report and in Snap's Transparency Reports. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Limited application. No category-specific Article 22 DSA trusted-flagger relationship has been identified for Fraud and Spam during the assessment period. Fraud- and spam-related notices submitted through any applicable priority-notice channels are processed under the procedures described in Section 5.10. Bad-faith or abusive reporting is addressed separately through the Article 23 DSA misuse-suspension framework described in Section 5.4.6.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap engages with EU and Member State financial-supervision frameworks where appropriate, [redacted]
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Harmful, False or Deceptive Practices category explainer, the Safety Center and Snap's Transparency Reports provide information on the substantive prohibitions, enforcement and prevalence, and the resources available to users. Snap's Transparency Reports also publish aggregate post-publication advertising reporting and removal volumes. The pre-publication advertising review and rejection figures relied upon in this assessment are recorded

[redacted]

	separately in the Likelihood analysis above. Further detail is provided in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. [redacted]
Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines apply without regard to whether fraud or spam content is user-generated or AI-generated, including in respect of AI-generated impersonation content and deepfake-driven scam content. Generative-AI features are subject to the content-authenticity measures described in Section 5.9, including watermarking and contextual identification. The 31 March 2025 Alan Turing Institute report on AI-enabled crime is recognised as identifying an emerging risk dimension in this category.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of Fraud and Spam on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture-including the absence of an open recommender feed and broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), and pre-distribution moderation-the proactive moderation suite for content-side Fraud and Spam, the pre-publication advertising review pipeline, the friending-protection measures relevant to fraud-related contact pathways, [redacted] supports the conclusion that the mitigation framework operates effectively against this risk. Between 1 July 2025 and 30 June 2026, [redacted] distinct advertisements completed final review and [redacted] advertisements were rejected across all advertising-policy reasons. The H2 2025 Transparency Report separately recorded 81,976 advertisements reported following publication and 47,266 removed. These figures, considered alongside the low H1 2026 VVR results, provide evidence concerning the operation of the advertising and content-moderation frameworks.

[redacted]

There is no change in the overall conclusion from the 2025 Report. Using the methodology retained for year-on-year comparability, content-side PVP decreased from [redacted] in April 2025 to [redacted] for the 90-day period ending 30 June 2026 and remains substantially below the [redacted] recorded in August 2023. The corresponding confirmed-violation PVP under the refined double-review methodology described in Section 6.4 was [redacted] The mitigation framework will continue to be monitored through the Section 6 risk-detection processes, with particular attention to emerging AI-enabled fraud patterns and developments arising from engagement concerning online financial misconduct.

4.1.12 Dissemination of Information Related to Other Illegal Activities

This Section 4.1.12 sets out Snap's assessment of the actual or foreseeable risk of harm arising from the dissemination of illegal content or activity on Snapchat's in-scope services that is not specifically addressed in Sections 4.1.1 through 4.1.11. This risk operates as a catch-all to ensure that the Category 1 systemic-risk assessment covers illegal content that may emerge in categories not anticipated at the time of this Report.

Harm Description

As a consequence of permitting users to publish content, it is possible that information related to illegal activities not specifically captured in the preceding categories may be disseminated on Snapchat. This section covers any such residual category of illegal content.

Although laws and regulations differ across the jurisdictions in which Snapchat operates, and Snapchat is a global community, the substantive prohibitions in the Community Guidelines (in particular, the Illegal or Regulated Activities category) prohibit, as a general matter, any activity that undermines public safety or violates human rights, the laws of the United States, or the laws of the country in which the user is located. The categories of conduct prohibited under this catch-all framework include, without limitation:

- the promotion of criminal activity not specifically addressed in the categories above (Sections 4.1.1-4.1.11);
- the facilitation or participation in cybercrime not specifically addressed elsewhere (including, where applicable, conduct adjacent to but distinct from the Fraud and Spam category in Section 4.1.11);
- any other content the dissemination of which constitutes a criminal offence under applicable EU or Member State law and that is not addressed by an existing Snap harm category.

The Article 16 DSA notice-and-action mechanism (Section 5.4.6) provides the principal operational pathway for receipt of notices alleging illegal content within this catch-all category.

[redacted]

The Article 9 / 10 DSA pathway provides the operational mechanism for receipt of orders from EU and Member State authorities to act against illegal content.

Likelihood

Internal data

The Article 16 DSA illegal content reporting option was introduced as a dedicated reporting pathway upon the entry into force of the DSA. Between 1 July 2025 and 30 June 2026, [redacted] reports were submitted through this pathway in the EU, compared with [redacted] reports during the preceding 12-month period. This represents an increase of [redacted] reports, or approximately [redacted]

No separate PVP or VVR is presented for this catch-all category because it does not correspond to a single defined policy or enforcement classification that can be sampled or measured consistently. Content falling within an established category is assessed in Sections 4.1.1-4.1.11. The likelihood assessment for this residual category therefore relies principally on Article 16 notice volumes and the Section 6 ongoing risk-detection processes.

Advertising and Lenses are assessed using separate pre-publication review, rejection, post-publication reporting and enforcement datasets. Because this section is a residual catch-all rather than a defined Advertising or Lens review category, no directly comparable category-specific rejection volume is presented.

The available data establishes the volume of reports submitted through the pathway but does not provide a reliable quantitative breakdown of report quality, actionability, resulting enforcement or the proportion of notices relating to conduct already addressed in Sections 4.1.1-4.1.11. Accordingly, the increase in reporting volume is treated as evidence of increased use of the pathway and is not, without further supporting information, treated as evidence of an increase in the prevalence of residual illegal content within this catch-all category.

External evidence

The Section 6 ongoing risk-detection processes include consultation with the Safety Advisory Board, engagement with trusted flaggers, monitoring through relevant industry groups and internal cross-functional working groups. The 2026 assessment identified non-coercive criminal recruitment of minors as a distinct risk pathway. Because that risk is characterised primarily as a minor-safety concern, the detailed assessment is recorded in Section 4.4.3, with a cross-reference from this catch-all section.

Those processes did not identify any other emerging illegal-activity category requiring a dedicated Category 1 risk for this Report. The Prevalence Testing and Transparency Reporting frameworks continue to monitor the established categories addressed in Sections 4.1.1-4.1.11.

[redacted]

Likelihood conclusion

On the basis of the available Article 16 reporting data and the Section 6 ongoing risk-detection inputs, which did not identify a distinct material category of residual illegal activity outside the risks addressed in Sections 4.1.1-4.1.11, the likelihood of dissemination of illegal content within this catch-all category on Snapchat's in-scope services is assessed at the **lowest likelihood category**. EU reporting volume through the Article 16 pathway increased from [redacted] reports during 1 July 2024-30 June 2025 to [redacted] reports during 1 July 2025-30 June 2026[64]. However, the available data does not identify the proportion of those reports that were actionable, resulted in enforcement or related specifically to this residual category rather than an existing Category 1 risk. The increase in report volume is therefore not treated as equivalent to an increase in the prevalence of residual illegal content. There is no change in the likelihood determination from the 2025 Report.

Severity

By definition, this catch-all category covers a potentially heterogeneous range of illegal content that is not addressed in Sections 4.1.1-4.1.11. The severity of an individual instance therefore depends on the nature of the underlying conduct, the persons affected and the scale and immediacy of the potential harm.

For overall risk-prioritisation purposes, the category is assessed as significant harm industry-wide. This reflects its residual nature and the fact that recurring categories of illegal content with established higher-harm profiles, including child sexual exploitation and abuse, sexual exploitation, terrorist content and specified illegal or regulated activities, are assessed separately in their dedicated sections.

The significant-harm classification is not a ceiling on the treatment of individual cases. Consistent with Section 3.5, any specific instance within this catch-all category that presents an imminent, credible threat to human life, safety or well-being is treated as a Level 1 overall risk and handled through the applicable prioritised-response and law-enforcement-escalation pathways.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this catch-all category.

[64] Internal Community Support data, EU reports submitted through the Article 16 DSA illegal-content reporting pathway, 1 July 2024-30 June 2026. Report volumes do not indicate actionability, resulting enforcement or applicable risk category.

[redacted]

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend illegal content, irrespective of the specific category in question; no interest categories supporting illegal content of any kind exist. The absence of an open recommender feed for unvetted user-generated content, the absence of a broad reshare functionality of Stories, and pre-distribution moderation for content reaching wide audiences reduce the structural opportunity for amplification of illegal content of any kind. Further detail in Section 5.6.
Content moderation systems	The proactive and reactive moderation framework described in Section 5.4 operates against illegal content of any kind, with the substantive scope of the Community Guidelines and the Article 16 DSA notice-and-action mechanism providing the operational pathways. The expansion of the illegal content reporting option in Section 5.4.6 supports the surfacing of illegal-content notices that fall outside the categories specifically addressed in Sections 4.1.1-4.1.11 of this Report.
Applicable terms and conditions and enforcement	The Snap Terms of Service, the Community Guidelines (in particular, the Illegal or Regulated Activities category), and the substantive prohibitions across all Community Guidelines categories together prohibit the use of Snapchat for any illegal activity. Where a notice through the Article 16 illegal content reporting option identifies content not addressed by an existing Community Guidelines category, Snap's Legal team - supported by external counsel as needed - reviews the notice to confirm illegality and to determine appropriate enforcement action. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising of illegal content of any kind. Pre-publication review applies to all advertising. Further detail in Section 5.7.
Data-related practices	Privacy and safety by design practices apply to product and feature changes that may bear on this risk. Changes are reviewed through the Online Interface Design Process described in Section 5.1.5 and the DDIA framework in Section 6.3.
Intentional manipulation (Article 34(2))	Bad-faith reporting through the Article 16 illegal content reporting option is addressed through the Article 23 DSA misuse-suspension framework described in Section 5.4.6. [redacted]

[redacted]

Amplification and rapid, wide dissemination	The platform-architecture factors that constrain amplification across all Category 1 categories (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation) apply equally to this catch-all category. Where any specific instance of an emerging illegal-content category presents an imminent credible threat to human life or to severe physical safety, the matter is escalated through the Content Crisis Response Protocol (Section 5.12) and through the prioritised-response and law-enforcement-escalation pathways described in Section 5.5.
Regional and linguistic aspects	Member State legal frameworks vary materially in respect of categories of illegal content (in particular, in respect of categories of content that are subject to specific Member State criminal codes). The Article 9 / 10 DSA pathway for receipt and processing of orders to act against illegal content from EU and Member State authorities provides the principal operational mechanism for Member State-specific illegal content not addressed by an existing Community Guidelines category. Engagement with Member State Digital Services Coordinators is maintained on a continuing basis.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, dissemination of content within this catch-all category is assessed at Level 3 overall potential risk prioritisation. This output corresponds to the matrix intersection of significant harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)-(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

[redacted]

Snapchat Design and Function - Art. 35(1)(a)	Applies. The platform-architecture factors that constrain amplification across all Category 1 categories apply equally to this catch-all category. The expansion of the illegal content reporting option (Section 5.4.6) elevates the visibility of the dedicated Article 16 DSA reporting pathway for illegal content not addressed by an existing Community Guidelines category. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Terms of Service and the Community Guidelines (in particular, the Illegal or Regulated Activities category) prohibit the use of Snapchat for any illegal activity. [redacted]. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. The Article 16 DSA notice-and-action mechanism (Section 5.4.6) provides the operational pathway for illegal content notices in this catch-all category. The elevation of the illegal content reporting option supports the visibility and accessibility of this pathway. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend illegal content of any kind; no corresponding interest categories exist. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit advertising of illegal content. Pre-publication review applies to all advertising. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. The Section 6 ongoing risk-detection processes, including consultation with the Safety Advisory Board, engagement with trusted flaggers, monitoring through industry groups and internal cross-functional working groups, are configured to identify emerging illegal-content categories that may warrant dedicated mitigation measures or a dedicated risk. The non-coercive criminal recruitment of minors is flagged as a distinct potential risk. Because that risk is characterised primarily as a minor-safety concern, the detailed assessment is recorded in Section 4.4.3 rather than as a separate Category 1 risk. No other emerging illegal-content category was identified as requiring a dedicated Category 1 for this Report. Further detail is provided in Section 6.

Trusted Flaggers - Art. 35(1)(g)	Applies. Trusted flaggers active across Community Guidelines categories may identify content falling within this catch-all category, and notices submitted by designated Article 22 DSA trusted flaggers are processed under the priority-notice procedures described in Section 5.10. Any recurring or emerging category identified through those notices is considered through the Section 6 ongoing risk-detection processes. Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. The Content Crisis Response Protocol (Section 5.12) operates across all illegal-content categories, including emerging categories not addressed by an existing Community Guidelines category. Cross-industry engagement is maintained on a continuing basis through industry groups including the [redacted] where relevant. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Illegal or Regulated Activities category explainer, the Safety Center, and Snap's Transparency Reports provide information on the substantive prohibitions applicable to this category. The Article 16 DSA illegal content reporting option is described in the Reporting and Notices documentation. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. The general minor-protection measures described in Section 5.8 apply to this category. Where a specific emerging category presents particular minor-protection concerns, the Section 6 ongoing risk-detection processes are configured to surface the concern for assessment.
Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines apply without regard to whether content in this catch-all category is user-generated or AI-generated. Generative-AI features are subject to the content-authenticity measures described in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors and the mitigations identified above, Snap has concluded that the mitigations in place for the dissemination of illegal content within this catch-all category on Snapchat's in-scope services are reasonable, proportionate and effective.

This conclusion is supported by the combined operation of the platform architecture, including the absence of an open recommender feed, short default retention for public stories (unless saved to Public Profile), and pre-distribution moderation; the Article 16 DSA notice-and-action mechanism; the Article 9 and 10 DSA pathway for receipt of orders from EU and Member State

[redacted]

authorities; Legal-team review of notices identifying potential illegal content not addressed by an existing Community Guidelines category; the Section 6 ongoing risk-detection processes; and the Content Crisis Response Protocol for elevated-risk events. EU reporting volume through the Article 16 pathway increased from [redacted] reports during 1 July 2024-30 June 2025 to [redacted] reports during 1 July 2025-30 June 2026. These figures demonstrate use of the pathway but do not, by themselves, establish the quality or actionability of the notices, the resulting enforcement outcomes or the proportion relating specifically to this residual catch-all category.

There is no change in the overall conclusion from the 2025 Report. This year's assessment identified non-coercive criminal recruitment of minors as a distinct potential risk pathway, which is assessed in Section 4.4.3 because it is characterised primarily as a minor-safety concern. The mitigation framework will continue to be monitored through the Section 6 risk-detection processes, with particular attention to notices submitted through the elevated illegal content reporting option and to any further emerging illegal-activity category that may warrant a dedicated risk.

4.2 Category 2 - Negative Effects on Fundamental EU Rights

This Section 4.2 sets out Snap's assessment of the actual or foreseeable negative effects of Snapchat's in-scope services on the exercise of fundamental rights protected by the Charter of Fundamental Rights of the European Union (the *Charter*). Consistent with Article 34(1)(b) DSA and Recital 81, the assessment covers, in particular, the rights to:

- human dignity (Article 1 Charter);
- freedom of expression and of information, including media freedom and pluralism (Article 11 Charter);
- respect for private and family life (Article 7 Charter);
- protection of personal data (Article 8 Charter);
- non-discrimination and freedom of thought, conscience and religion (Articles 21 and 10 Charter);
- the rights of the child (Article 24 Charter); and
- consumer protection (Article 38 Charter).

The right to property (Article 17 Charter), insofar as it concerns intellectual property on Snapchat's in-scope services, is addressed substantively in Section 4.1.5 and cross-referenced in Section 4.2.8.

Consistent with Recital 81, the Category 2 assessment considers, in particular, risks arising from (i) the design of Snap's algorithmic and interface systems and (ii) misuse of the service through the submission of abusive notices or other methods for silencing speech or hampering competition. In respect of the rights of the child, the assessment further considers how easy it is for minors to

[redacted]

understand the design and functioning of the service, and how minors can be exposed through the service to content that may impair their health, physical, mental or moral development - including through interface design that may, intentionally or unintentionally, exploit the weaknesses and inexperience of minors or contribute to addictive behaviour.

Category 2

Consistent with the methodology in Section 3, the Category 2 assessment in this Report includes the following refinements:

- **Default settings for minor accounts.** Default settings for minor accounts - including push-notification, Find Friends visibility and contact-permission defaults - are analysed as a design consideration bearing on the rights to private life (Section 4.2.3), data protection (Section 4.2.4) and the rights of the child (Section 4.2.6).
- **Minor-account onboarding flow.** The account-creation flow presented to users creating minor accounts is analysed as a distinct design consideration under the rights of the child (Section 4.2.6), addressing the order and prominence of feature-activation and data-sharing prompts relative to safety features and settings.
- **Discoverability of Family Center, Privacy Center and Safety Center.** The discoverability and accessibility of Family Center, Privacy Center and Safety Center - [redacted] - are analysed as a design consideration bearing on the rights of the child (Section 4.2.6) and on the transparency dimension of Articles 34 and 35 DSA.
- **Ease of notifying potentially illegal content.** The ease with which users can submit notices concerning content they believe to be illegal - and the parity of that pathway with other reporting pathways - is analysed as a design consideration bearing on freedom of expression and information (Section 4.2.2), alongside its treatment under Article 16 DSA.

Where these refinements bear on a specific Category 2 risk, they are reflected in the relevant sections 4.2.1 through 4.2.8. As a general matter, they do not alter the prioritization determinations reached in prior Reports, but they are intended to support a transparent Article 35(1) proportionality analysis.

Summary of Category 2 determinations

The table below summarises Snap's assessment of likelihood, harm classification, overall potential risk prioritization, and the conclusion on mitigations under Article 35(1) DSA for each Category 2 risk. Where a specific instance presents an imminent, credible risk of severe harm, including a threat to human life, safety or well-being, the case-by-case framework in Section 3.5 applies and the instance is treated as Level 1, irrespective of the category-level prioritization shown below.

Risk	Likelihood	Harm classification	Risk prioritization	Conclusion (Art. 35(1) DSA)
4.2.1 Right to Human Dignity	Lowest	Severe	**Level 1***	Reasonable, proportionate and effective
4.2.2 Freedom of Expression and Information	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective
4.2.3 Right to Private Life	Lowest	Serious	**Level 3**	Reasonable, proportionate and effective
4.2.4 Right to Data Protection	Highest	Severe	**Level 1**	Reasonable, proportionate and effective
4.2.5 Non-Discrimination and Freedom of Religion	Lowest	Serious	**Level 3**	Reasonable, proportionate and effective
4.2.6 Rights of the Child	Lowest	Severe	**Level 1***	Reasonable, proportionate and effective; ongoing engagement on EU-wide guidance
4.2.7 Consumer Protection	Lowest	Significant	**Level 3**	Reasonable, proportionate and effective
4.2.8 Right to Property (IP)	—	—	**—**	Addressed in Section 4.1.5 (IP infringement); not repeated

**Case-by-case upward adjustment from the default Section 3.5 matrix output; see Section 4.2.6 for the rationale.*

Note on the Right to Data Protection

The right to data protection remains the only risk in this Report assessed at the highest likelihood category. This classification reflects the nature of the right - which, as a general matter, is engaged by every instance of personal data processing on the service, rather than solely by content that violates Snap's terms and conditions - and is not an indicator of weaker mitigation. The detailed analysis, including the engagement of data-protection risk with the default-settings considerations identified above, is provided in Section 4.2.4.

Each Category 2 risk is addressed in turn in the sub-sections that follow.

[redacted]

4.2.1 Right to Human Dignity

This Section 4.2.1 sets out Snap's assessment of the actual or foreseeable negative effects of Snapchat's in-scope services on the Right to Human Dignity.

Snapchat's in-scope public surfaces, including Spotlight, Discover, Place Stories on Snap Map, Public Profiles, Lenses and Advertising, carry the potential for the dissemination or presentation of content that may undermine Human Dignity. Without mitigation, such content could include, without limitation:

- human trafficking and related coercion for sexual exploitation;
- child sexual abuse material;
- terrorist content;
- content that glorifies or promotes self-harm, including self-injury, suicide or eating disorders; and
- incitement to violence or hatred directed against a group of persons or a member of a group on any of the grounds referred to in Article 21 of the Charter.

Likelihood

The likelihood analysis for the Right to Human Dignity is derived from the likelihood determinations for the underlying Category 1 risks that bear on this right. The current determinations, drawn from Section 4.1 of this Report, are:

Underlying Category 1 risk	likelihood
4.1.1 Dissemination of CSEA Content	**Lowest**
4.1.2 Dissemination of Illegal Hate Speech	**Lowest**
4.1.4 Dissemination of Terrorist Content	**Lowest**
4.1.6 Dissemination of Adult Sexual Content	**Lowest**
4.1.8 Self-Harm, Suicide and Eating Disorders	**Lowest**

No separate PVP or VVR is presented for the Right to Human Dignity because this is a rights-based aggregation of underlying content risks rather than a standalone policy or enforcement category. Relevant PVP and VVR evidence is presented in the corresponding Category 1 sections: 4.1.1, 4.1.2, 4.1.4, 4.1.6 and 4.1.8.

External Evidence

External evidence supports the foreseeability of dignity-related harms across online services, while not providing a Snapchat-specific prevalence estimate. Research by the European Union

Agency for Fundamental Rights found that abusive comments, harassment and incitement to violence can remain accessible after platform content-moderation systems have been applied, with women and certain minority groups among the principal targets.[65] This evidence is used as contextual support for the rights pathway assessed here; the relative-likelihood determination remains based on the PVP, VVR and enforcement evidence presented in the underlying Category 1 sections.

On the basis of Snapchat's design factors, the available procedural safeguards, the lowest-likelihood determinations for the underlying Category 1 risks, and the external evidence described above, the likelihood of a negative effect on the Right to Human Dignity on Snapchat's in-scope services is assessed at the **lowest likelihood category**.

Severity

The Council of Europe has recognised that misuse of social media can trigger harmful consequences threatening Human Dignity, citing examples that include hate speech, incitement to violence and discrimination.[66] The severity analysis for this right aggregates the severity determinations for the underlying Category 1 risks, drawn from Section 4.1 of this Report:

Underlying Category 1 risk	Harm classification
4.1.1 Dissemination of CSEA Content	**Severe**
4.1.2 Dissemination of Illegal Hate Speech	**Significant**
4.1.4 Dissemination of Terrorist Content	**Serious**
4.1.6 Adult Sexual Content - sexual crimes	**Serious**
4.1.6 Adult Sexual Content - other	**Significant**
4.1.8 Self-Harm, Suicide and Eating Disorders	**Serious**

Consistent with prior Reports, the Right to Human Dignity is assessed at the severity of the most severe underlying risk: **severe harm industry-wide**. Certain underlying harms, particularly CSEA and human trafficking, may constitute severe harms and warrant heightened enforcement scrutiny.

DSA Risk Factors

[65] European Union Agency for Fundamental Rights, Online Content Moderation – Current Challenges in Detecting Hate Speech (29 November 2023).
[66] Council of Europe, *'Social media: social threads or threats to human rights'* January 2017, url.

[redacted]

The service and general risk factors set out in Article 34(2) DSA have been considered in the context of the Right to Human Dignity. Because these factors, as a general matter, operate similarly across the underlying Category 1 harms and are addressed through horizontal mitigation measures that cut across multiple systemic risks, the detailed factor-by-factor analysis is not repeated in this section. The reader is directed to the Article 34(2) analyses in the corresponding Category 1 Sections: Section 4.1.1 (CSEA Content), Section 4.1.2 (Illegal Hate Speech), Section 4.1.4 (Terrorist Content), Section 4.1.6 (Adult Sexual Content) and Section 4.1.8 (Self-Harm, Suicide and Eating Disorders).

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5 and the case-by-case departure permitted by the methodology, the Right to Human Dignity is assessed at Level 1 overall potential risk prioritisation. Although the likelihood is at the lowest relative level and the default matrix output would ordinarily be Level 3, the severity of the underlying harms-including CSEA, terrorist content and human trafficking-warrants a departure to Level 1. This is consistent with the methodology set out in Section 3.5 and with the approach taken in prior Reports. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA in the table below. Because the mitigation set for the Right to Human Dignity is substantially co-extensive with the mitigations for the underlying Category 1 risks, the descriptions are illustrative; detailed descriptions of each mitigation are provided in Section 5, cross-referenced by sub-section.

[redacted]

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to include proactive moderation for CSEA and other illegal content bearing on Human Dignity. In-app reporting tools are provided for users to report such content to the Trust & Safety team. Where the Trust & Safety team identifies a Snapchatter in distress, support resources (including the Here For You tool for suicide-related search terms) and, where appropriate, emergency-personnel notification pathways are available. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines prohibit CSEA and other illegal content that undermines Human Dignity, and are enforced in accordance with the severity of the harm. Further detail in Section 5.2.
Moderation - Art. 35(1)(c)	Applies. Proactive and reactive moderation procedures address CSEA, hate speech, terrorist content, adult sexual content, and content glorifying self-harm, suicide or eating disorders. [redacted]
Algorithmic Systems - Art. 35(1)(d)	Applies. Our algorithmic systems do not knowingly recommend content that would negatively affect the Right to Human Dignity i.e. there are no interest categories that we consider to negatively affect Human Dignity; no corresponding interest categories exist for such content. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising policies prohibit content that undermines Human Dignity, including hate speech and discriminatory content. The mitigations identified elsewhere in this table apply to advertising systems where relevant. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing, Transparency Report outputs and enforcement data are used to monitor risks relevant to Human Dignity on a continuous basis. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Cooperation with trusted flaggers is established for CSEA and other illegal content that undermines Human Dignity. Further detail in Section 5.10.

[redacted]

Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Participation in industry coordination groups includes the [redacted] among others. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. Guidance on terms, harms, moderation and enforcement is published in the Safety Center, with in-app support surfaces such as Here For You for sensitive search terms. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. [redacted] Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Partially applies. Measures are in place to mitigate (i) the use of generative-AI tools within the service to create illegal or otherwise violating content, and (ii) the dissemination on in-scope services of illegal or otherwise violating content created using generative-AI tools on other platforms. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the Right to Human Dignity on Snapchat's in-scope services are reasonable, proportionate and effective.

4.2.2 Right to Freedom of Expression and Assembly

This Section 4.2.2 sets out Snap's assessment of the actual or foreseeable negative effects of Snapchat's in-scope services on the rights to freedom of expression and assembly under the EU Charter of Fundamental Rights, the European Convention on Human Rights, and the corresponding international instruments.

Snapchat's in-scope services are intended, among other purposes, to enable users to express themselves, share visual content, and connect with others. As a general matter, the services support the exercise of the rights to expression and assembly. At the same time, the operation of any digital content service can give rise to actual or foreseeable negative effects on these rights, and the assessment recorded here addresses those negative-effect pathways.

Risk pathways

The following risk pathways are considered in this section:

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]

Recent developments

The following developments have been identified through Snap's ongoing risk assessment process as bearing on the freedom of expression and assembly risk pathways set out above.

- [redacted]

[67] Council of Europe, Guidance Note on Content Moderation: Best Practices towards Effective Legal and Procedural Frameworks for Self-Regulatory and Co-Regulatory Mechanisms of Content Moderation (2021), identifying proportionality, safeguards against over-compliance and discrimination, and independent review mechanisms as principles of a human-rights-based approach to content moderation.

[redacted]

- [redacted]

Likelihood

Algorithmic suppression, over-moderation and self-censorship are difficult to measure directly. Snap maintains user-feedback, community-support, internal complaint-handling and appeal pathways through which users can raise concerns regarding moderation or distribution decisions. These mechanisms also provide signals that may support the identification of recurring or systemic issues affecting lawful expression.

The design of Snapchat's in-scope services limits the scale at which an erroneous moderation, classification or ranking decision could affect lawful expression. Public-content recommendation operates only on defined surfaces and subject to applicable recommendation-eligibility requirements, while users retain access to internal complaint-handling and out-of-court dispute-settlement mechanisms where content is restricted or removed.

No quantitative conclusion is drawn from internal complaint or reversal volumes for this assessment period. No separate PVP or VVR is presented because freedom of expression and assembly is a rights-based risk concerning moderation, ranking, reporting and redress processes rather than a standalone content-policy or enforcement category. The assessment therefore relies on the operation of internal complaint-handling, appeal, out-of-court dispute-settlement, community-support and misuse-detection pathways. Advertising and Lenses are assessed through their applicable review, reporting and enforcement frameworks rather than through a freedom-of-expression-specific PVP or VVR measure; no directly comparable category-specific rejection volume is presented.

External Evidence

External evidence confirms that erroneous or disproportionate content-moderation decisions are a foreseeable industry-wide risk and that effective redress mechanisms can materially affect moderation outcomes. In February 2026, the European Commission reported that users had appealed more than 165 million content-moderation decisions through platforms' internal complaint-handling mechanisms since the DSA became applicable, with almost 30% of those decisions reversed.[68] These cross-platform figures are not specific to Snapchat and do not provide a measure of the prevalence of erroneous moderation on Snapchat's in-scope services.

[68] European Commission, 'Two Years of Digital Services Act Allows 50 Million Content Moderation Decisions by Platforms to Be Reversed' (17 February 2026).

They nevertheless support the relevance of the internal complaint-handling, statement-of-reasons and out-of-court dispute-settlement safeguards considered in this assessment.

On the basis of Snapchat's design factors, the available procedural and redress safeguards, and the external evidence described above—recognising that the external evidence establishes an industry-wide risk rather than a Snapchat-specific incidence rate—the likelihood of a negative effect on freedom of expression and assembly on Snapchat's in-scope services is assessed at the **lowest likelihood category**.

Severity

Freedom of expression and assembly are foundational rights protected under the EU Charter, the ECHR, and the corresponding international instruments. The Council of Europe has recorded that social-media services are key actors in the regulation of information flow online, and that the way in which they operate has a significant impact on freedom of expression, including freedom of information.[69]

Industry-level risks bearing on severity include: algorithmic suppression patterns that reduce visibility of speech from minority or marginalised communities; automated-moderation systems that over-remove ambiguous or borderline content; coordinated reporting campaigns directed at specific creators; and undue restriction of advertising content. These risks have the potential, in the aggregate and over time, to produce material harm to public discourse.

In the specific context of Snapchat's in-scope services, the severity of negative effects is, however, more limited than at services whose primary architecture is the broadcast of unvetted user-generated content to a public audience. Snapchat is, as a general matter, not a service whose primary use case is the public dissemination of political or activist content; the in-scope services do not, in the ordinary course, present an open recommender feed for unvetted user-generated content; and the structural amplification of speech is bounded by the design choices recorded in Section 5.1. On that basis, the severity of negative effects on freedom of expression and assembly is assessed as **significant harm industry-wide**, rather than serious or severe. There is no change in this determination from the 2025 Report.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services. While the factors could, in principle, influence this systemic risk, their overall impact is limited in view of the mitigations described in Snap's Mitigations below and in Section 5.

[69] Council of Europe, *'Social media: social threads or threats to human rights'* January 2017, url.

[redacted]

Design of recommender systems and other algorithmic systems	[redacted] Further detail is provided in Section 5.6.
Content moderation systems	Proactive and reactive moderation procedures operate to limit the visibility of unvetted content reaching wide audiences on in-scope services. [redacted] Where moderation outcomes are challenged, the internal complaint-handling pathway under Article 20 DSA and the out-of-court dispute settlement pathway under Article 21 DSA are available to affected users. Further detail in Section 5.4 (Content Moderation) and Section 5.5 (Enforcement).
Applicable terms and conditions and enforcement	The Community Guidelines define the prohibited categories of content (Sexual Content; Harassment and Bullying; Threats, Violence and Harm; Harmful False or Deceptive Information; Illegal or Regulated Activities; Hateful Content, Terrorism and Violent Extremism). Enforcement is conducted in accordance with the Terms and consistent with the Article 16 notice-and-action and Article 17 statement-of-reasons obligations, as further described in Section 5.5. Further detail in Section 5.2 (Terms) and Section 5.5 (Enforcement).
Systems for selecting and presenting advertisements	Advertisements are subject to pre-publication review against the Advertising Policies and the Community Guidelines, with reactive reporting pathways available. As of October 2025, and consistent with the EU's Political Advertising Regulation, Snap has discontinued political advertising in the European Union; this measure is recorded in Section 4.3.1 (Negative Effects on Democratic and Electoral Processes) and Section 5.7 (Advertising Systems). The discontinuation of political advertising is a regulatory-compliance measure rather than an editorial intervention into political speech. Further detail in Section 5.7.

[redacted]

Data-related practices	Privacy and safety by design practices, documented through the PASS review and the DDIA framework, apply to product and feature changes that may bear on this risk. [redacted] in Section 5.1.5 and, where applicable, through the DDIA framework described in Section 6.3.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	In-scope services are configured in a way that limits both the amplification of harmful content and the structural opportunity for over-suppression of lawful speech. Short default retention for public stories (unless saved to Public Profile), the absence of a broadcast feed for user-generated political content, and the pre-distribution moderation [redacted] combine to keep the distribution architecture narrow. Lawful expression remains possible across the full range of in-scope services, subject to the Community Guidelines and the Content Guidelines for Recommendation Eligibility.
Regional and linguistic aspects	The service is available in multiple EU languages, the Terms are published in all official EU languages, and moderators operate across multiple EU languages. Moderation guidance is calibrated to Member State legal contexts where required (for example, in respect of categories of content that are illegal in specific Member States but not others).

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, negative effects on freedom of expression and assembly are assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of significant harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

	Lowest likelihood	Medium likelihood	Highest likelihood

[redacted]

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. Foundational design choices reduce both the risk of harmful content reaching wide audiences (limiting harm to other rights) and the structural opportunity for over-suppression of lawful speech: short default retention for public stories (unless saved to Public Profile); bidirectional friending; the absence of a broadcast feed for user-generated political content; pre-distribution moderation for content reaching wide audiences. The minor-account default settings recorded in Section 5.1.3 are calibrated to be necessary, proportional and legitimate (see Section 6.1), recognising that minor users retain rights to expression and information that are protected from over-restriction. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines and the Content Guidelines for Recommendation Eligibility define prohibited content with reference to substantive harms rather than open-ended categories. Where an enforcement decision is taken, an Article 17 statement of reasons is provided and the Article 20 internal complaint-handling pathway is available. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. [redacted] Further detail in Section 5.4.

[redacted]

Algorithmic Systems - Art. 35(1)(d)	Applies. Recommender systems are not configured to suppress or deprioritise otherwise lawful content on the basis that it expresses a particular viewpoint or relates to lawful assembly. [redacted] Users in the European Union can disable personalised public-content recommendations through the "European Union Controls" settings described in Section 5.6. Further detail is provided in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising review is conducted against the Advertising Policies and the Community Guidelines. As of October 2025, Snap has discontinued political advertising in the European Union, consistent with the EU's Political Advertising Regulation. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Internal complaint-handling and appeal outcomes, enforcement reversals, out-of-court dispute-settlement outcomes, Article 23 misuse-suspension activity, and the volume and nature of community-support requests are monitored to identify trends that may bear on this risk. Transparency Report outputs and relevant enforcement data provide additional operational context. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA are processed through the applicable priority-notice procedures. Bad-faith or manifestly unfounded notices, including notices intended to suppress lawful expression or assembly, are addressed separately through the Article 23 DSA misuse-suspension framework described in Section 5.4.6. Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Engagement with external experts on platform governance, freedom of expression and content moderation includes Snap's Safety Advisory Board and individual collaborations with experts identified in Section 6.1.2 [redacted]. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Privacy, Safety and Policy Hub publishes the Community Guidelines, the Content Guidelines for Recommendation Eligibility, the Advertising Policies and the supporting explainers. Statements of reasons under Article 17 DSA are issued in connection with enforcement decisions. Further detail in Section 5.3.

[redacted]

Protection of Minors - Art. 35(1)(j)	Applies. Mitigations for minors have been assessed against the rights to expression and information of minor users, in line with Article 12 and Article 17 of the UN Convention on the Rights of the Child. [redacted] Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Partially applies. Content-authenticity measures address the dissemination of deceptive synthetic content, which can itself bear on the integrity of public discourse. Watermarking and contextual identification of generative-AI content are designed to inform recipients without categorically restricting access to such content. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for negative effects on freedom of expression and assembly on Snapchat's in-scope services are reasonable, proportionate and effective.

The proportionality framing is significant for this assessment: the mitigations applied to other systemic risks (in particular, content moderation, advertising restriction, and minor-account design) operate in tension with freedom of expression and assembly. Snap's approach - documented through the Platform Principles-based Framework in Section 6.1 and the 2026 methodology refinement on Article 35(1) proportionality - is to apply the lightest mitigation that is necessary, proportional and legitimate for the underlying risk, in order to avoid disproportionate restriction of expression.

There is no change in the overall conclusion from the 2025 Report.

4.2.3 Right to Private Life

This Section 4.2.3 sets out Snap's assessment of the actual or foreseeable risk of negative effects on the Right to Private Life arising from the operation of Snapchat's in-scope services. The Right to Private Life is recognised under Article 7 of the EU Charter (respect for private and family life, home and communications) and Article 8 of the European Convention on Human Rights, and is closely related to (but distinct from) the Right to the Protection of Personal Data addressed in Section 4.2.4.

Online platforms can be used to spread content that undermines respect for private and family life. On Snapchat, content of this character could conceivably appear in any of the in-scope

[redacted]

services displaying user-generated content, including videos featured on Spotlight and Discover, content displayed on Snap Map, content featured on Public Profiles, and content carried in Lenses or advertisements. This Section assesses the risk to the Right to Private Life given the operational mitigation framework in place across the in-scope services.

Category 1 Risk Relevance

The Right to Private Life is the fundamental-rights mirror of several Category 1 risks where the operational mitigation framework directly bears on this risk: in particular, Section 4.1.6 (Adult Sexual Content, including NCII), Section 4.1.7 (Harassment and Bullying, including doxxing), and Section 4.1.10 (Harmful False Information, including impersonation). The mitigation refinements (minor-account default settings and onboarding) reduce minor exposure to privacy-violating contact and content as a secondary effect; the friending-protection refinements (Section 5.1.3 / 5.4.4) reduce the contact pathway for privacy-violating conduct directed at minors; and the Article 16 reporting-pathway and Article 23 misuse-suspension changes recorded in Section 5.4.6 support the reactive-reporting dimension of the framework.

Likelihood

Three internal data sources are tracked in respect of the Right to Private Life:

- **Prevalence Testing - invasion-of-privacy category.** This category captures Prevalence Testing outputs in respect of content that, on assessment, would be classified as invasive of privacy.
- **Privacy operations requests.** Annual volume of operational requests submitted through the dedicated privacy-request channels (covering data subject rights requests, account-related privacy concerns, and similar).
- **Data Protection Officer queries flagged as relating to GDPR or Member State privacy laws.** Annual volume of queries directed to Snap's Data Protection Officer that have been flagged at intake as relating to GDPR or Member State privacy laws.

The trends across the recent assessment cycles are recorded below:



[redacted]

The data shows: (i) a marked decrease in invasion-of-privacy PVP from [redacted] in the 2025 Report to [redacted] for the 90-day period ending 30 June 2026 under the methodology retained for year-on-year comparability in Section 6.4; (ii) a further decrease in annual Privacy Operations requests from [redacted] and (iii) an increase in DPO queries flagged as relating to GDPR or Member State privacy laws, from [redacted] during the 2026 assessment period. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation invasion-of-privacy PVP for the 90-day period ending 30 June 2026 was [redacted] The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction. [redacted]

Generative-AI tools may also make it easier for users to draft formal privacy-related correspondence and identify relevant contact addresses. Accordingly, the increase in DPO-query volume is treated as an operational intake trend and is not, on its own, interpreted as evidence of a corresponding increase in the incidence of privacy harm on Snapchat.

No separate VVR is presented for the Right to Private Life because this rights-based assessment draws on a broader set of indicators, including invasion-of-privacy PVP, privacy operations requests and DPO queries. Relevant VVR evidence is addressed, where available, in the corresponding Category 1 Sections, including Sections 4.1.6, 4.1.7 and 4.1.10. Advertising and Lenses are assessed through separate pre-publication review, reporting and enforcement frameworks; no directly comparable privacy-specific rejection volume is presented.

Likelihood conclusion

On the basis of the internal data recorded above, including the decrease in invasion-of-privacy PVP from [redacted] to [redacted] under the methodology retained for year-on-year comparability, the supplementary confirmed-violation PVP of [redacted] under the refined double-review methodology described in Section 6.4, and the further decrease in Privacy Operations request volumes, the likelihood of the in-scope services having a negative effect on the Right to Private Life is assessed at the **lowest likelihood category.** The increase in DPO queries during the assessment period has also been considered; however, given the intake-based methodology and the changes in user contact behaviour described above, that metric is not treated as a direct measure of the incidence or prevalence of privacy harm. There is no change in this determination from the 2025 Report.

Severity

Online platforms can be used to spread content that undermines respect for private and family life, home and communications, and personal data, including by:

[redacted]

- distributing surreptitious or non-consensual imagery of intimate activities, private movements or identifying information;
- distributing deceptive content involving the impersonation of friends, public figures or celebrities for harmful (rather than satirical) purposes;
- depicting Teens without parental consent;
- aggregating or revealing private information for the purpose of harassment (doxxing).

Without mitigations, content that undermines the Right to Private Life can find reach, particularly where it is of high public interest (such as celebrity voyeurism or doxxing of named persons) or where it appeals to prurient interests (for example, certain non-consensual intimate imagery). The severity of the risk lies not only in the immediacy of harm but also in the potential for enduring psychological, reputational and social damage. Victims may experience anxiety, fear, social isolation, or ongoing harassment, particularly where content is widely circulated or persists in archived form. In the case of minors, the impact can be particularly profound, including effects on development, education and digital confidence.

Recognising both the immediate and the enduring dimensions of the harm, the severity of negative effects on the Right to Private Life is assessed as **serious harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are designed not to knowingly recommend content that impedes the exercise of the Right to Private Life, including content depicting individuals without their consent, doxxing content, or impersonation content. The absence of an open recommender feed for unvetted user-generated content for unvetted user-generated content, the absence of a broad reshare functionality of Stories, and pre-distribution moderation for content reaching wide audiences reduce the structural opportunity for amplification of privacy-violating content. Further detail in Section 5.6.
Content moderation systems	Proactive and reactive moderation procedures operate to detect and remove content that impedes the exercise of the Right to Private Life. [redacted] Reactive moderation is available through the in-app reporting flow, including the dedicated reporting reasons

[redacted]

	for harassment, sexual content, and other privacy-related categories. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines prohibit impersonation; the prohibition on doxxing under the Harassment and Bullying category protects against the unauthorised disclosure of private information; the substantive prohibitions on NCII under the Sexual Content category (Section 4.1.6) protect against non-consensual intimate imagery; and the Commercial Content Policy prohibits non-consensual sexual material. The Spotlight Terms require users to have any necessary third-party rights, including music copyrights and rights of publicity, for content submitted to Spotlight. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising that displays information violating the law or causing certain privacy-related harms. A mix of pre-publication automated and human review applies to all advertising. User-facing controls allow advertisements to be hidden or reported. Further detail in Section 5.7.
Data-related practices	[redacted]
Intentional manipulation (Article 34(2))	[redacted]

[redacted]

Amplification and rapid, wide dissemination	Amplification of privacy-violating content on in-scope services is structurally constrained: Snapchat is not, as a general matter, an attractive distribution platform for content of this type (the absence of an open recommender feed, the absence of a broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), and the bidirectional friending requirement together limit the audience for such content); content reaching wide audiences on Spotlight and Discover is moderated using a combination of automated and human moderation prior to wide distribution; and Lenses and advertisements are subject to pre-publication review. [redacted]
Regional and linguistic aspects	Member State-specific privacy frameworks (in particular, GDPR-implementing legislation and any Member State-specific civil-privacy frameworks) are reflected in the substantive moderation framework where required. The Article 16 DSA notice-and-action mechanism, the dedicated privacy reporting pathways, and the in-app reporting flow are available in multiple EU languages.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, the risk to the Right to Private Life is assessed at Level 3 overall potential risk prioritization. This output corresponds to the matrix intersection of serious harm industry-wide with lowest likelihood, without case-by-case deviation. It should be noted that this is a decrease as Right to Private Life was assessed as a Level 2 overall potential risk prioritization in the 2025 Risk Assessment.



Snap's Mitigations

[redacted]

The mitigations in place for this risk are mapped to Article 35(1)(a)-(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. Privacy is one of Snap's foundational platform-governance values, reflected in the architectural design of Snapchat. Specific design features addressing this risk include: short default retention for public stories (unless saved to Public Profile); the bidirectional friending requirement, which limits unsolicited contact; private friend lists by default; recommended content in public feed is moderated; no broad reshare functionality of Stories; pre-distribution moderation for content reaching wide audiences. Lenses are designed to use object detection rather than facial identification. Snap Map operates under privacy-by-default principles, with location sharing optional and granular. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Harassment and Bullying prohibit doxxing (the sharing of another person's private information without their knowledge and consent) and the sharing of images of a person in a private space without their knowledge and consent. The Community Guidelines on Sexual Content prohibit non-consensual intimate imagery (NCII). The Community Guidelines prohibit impersonation. The Commercial Content Policy prohibits non-consensual sexual material. The Spotlight Terms require users to have any necessary third-party rights for content submitted to Spotlight. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. [redacted]
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend privacy-violating content; no corresponding interest categories exist. [redacted] Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit advertising that displays information violating the law or causing certain privacy-related harms. Pre-publication review applies to all advertising. Further detail in Section 5.7.

[redacted]

Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Privacy-related metrics are monitored on a continuing basis through Prevalence Testing (invasion-of-privacy category), through the volume and nature of privacy operations requests, and through the volume and nature of queries to Snap's Data Protection Officer. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Trusted-flagger relationships in adjacent categories (NCII in Section 4.1.6, harassment in Section 4.1.7) also support the privacy framework.
Codes and Crisis Protocols - Art. 35(1)(h)	Limited application. No specific code of conduct is in operation in respect of the Right to Private Life on the in-scope services. Snap engages with privacy-rights organisations, civil-society groups and human-rights experts on an ongoing basis to inform approach. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Privacy Policy, the Safety Center, the Privacy Center, and Snap's Transparency Reports provide information on the substantive privacy prohibitions, on user controls, and on the resources available to users encountering privacy violations. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. [redacted] Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. The Community Guidelines apply without regard to whether privacy-violating content is user-generated or AI-generated, including in respect of synthetic content depicting real persons (deepfakes, AI-manipulated impersonation content, synthetic NCII). The content-authenticity measures described in Section 5.9 (watermarking and contextual identification of generative-AI content) operate alongside the substantive prohibitions. Generative-AI Lens flows include input/output scanning to mitigate the misuse pathway.

[redacted]

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the protection of the Right to Private Life on Snapchat's in-scope services are reasonable, proportionate and effective.

Privacy is one of Snap's foundational platform-governance values and the architectural design of Snapchat is consistent with that value: short default retention for public stories (unless saved to Public Profile); the bidirectional friending requirement; private friend lists; recommended content in public feed is moderated; no broad reshare functionality of Stories; [redacted] and the use of object detection rather than facial identification in Lenses. The substantive prohibitions across the Community Guidelines (in particular, the prohibitions on doxxing, impersonation and NCII, addressed in Sections 4.1.6, 4.1.7 and 4.1.10) provide the operational pathway for privacy-related enforcement. [redacted]. Specific minor-protection measures [redacted] provide elevated protection for the most vulnerable users.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, including monitoring of invasion-of-privacy PVP, privacy operations requests and DPO GDPR/Member State queries.

4.2.4 Right to the Protection of Personal Data

This Section 4.2.4 sets out Snap's assessment of the actual or foreseeable risk of negative effects on the Right to the Protection of Personal Data arising from the operation of Snapchat's in-scope services. The Right to the Protection of Personal Data is recognised under Article 8 of the EU Charter, given operational specificity by the General Data Protection Regulation (GDPR), and is closely related to (but distinct from) the Right to Private Life addressed in Section 4.2.3.

This risk receives the most elevated risk prioritisation among the Category 2 fundamental-rights risks (Level 1), reflecting the structural fact that Snap, as a designated VLOP under the DSA, processes a substantial volume of personal data of EU/EEA users in the course of operating Snapchat's in-scope services. The likelihood determination for this risk is driven not by the empirical incidence of complaints or violations (which is low in absolute terms, as recorded below) but by the structural exposure inherent in the volume of processing.

GDPR Framework

[redacted]

As part of this assessment we have further enhanced the Right to the Protection of Personal Data in two respects: [redacted]

Likelihood

As recorded in Section 4.2.3 (Right to Private Life), the absolute volume of privacy-related queries from EU users is low relative to the population of EU monthly active recipients on Snapchat. The number of privacy operations requests in the EU and number of queries to the Data Protection Officer flagged as relating to GDPR or Member State privacy laws have been set out below:

[redacted]

Snap has an average of approximately 96.5 million monthly active recipients in the EU during the 2026 assessment period.. Against that user-base scale, the absolute volume of privacy-related queries is materially low. The Prevalence Testing data for the related “invasion of privacy” content category similarly shows very low absolute prevalence. Using the methodology retained for year-on-year comparability in Section 6.4, invasion-of-privacy PVP decreased from [redacted] in the 2025 Report to [redacted] for the 90-day period ending 30 June 2026. Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP was [redacted] The refined figure is presented as supplementary 2026 context and is not used to determine the year-on-year direction.

No separate VVR is presented for the Right to the Protection of Personal Data because this is a rights-based and structurally driven risk rather than a standalone content-policy or enforcement category. The invasion-of-privacy PVP figure is used as a related content-risk indicator and

[redacted]

should not be interpreted as a direct measure of compliance with data-protection requirements. Advertising, Lenses, Discover and the other in-scope services are assessed through the product- and processing-specific safeguards described below, including PASS Review, DPIA and DDIA scrutiny, advertising, targeting controls, data minimisation and on-device Lens processing-rather than through a data-protection-specific PVP, VVR or rejection-volume measure.

Notwithstanding the low absolute volume of privacy-related queries, a substantial volume of personal data is processed by Snap in the operation of Snapchat's in-scope services. [redacted]

The volume and continuity of this processing constitute a structural exposure that, in the absence of effective safeguards, could give rise to data-protection harm.

Consistent with prior Reports, this structural exposure is the principal driver of the likelihood determination for this risk. The empirical observations recorded above support the determination that the safeguards in place have, in operation, kept the user-facing incidence of data-protection harm low; but the underlying structural exposure remains, by virtue of the platform model itself.

On the basis of the structural-exposure rationale recorded above, driven by the volume and continuity of personal data processing inherent in the operation of the in-scope services, rather than by the empirical incidence of complaints (which is low), the likelihood of negative effects on the Right to the Protection of Personal Data is assessed at the **highest likelihood** category. There is no change in this determination from the 2025 Report. This determination is structural rather than incident-driven and remains stable across assessment cycles unless and until the underlying processing footprint or the safeguards regime materially changes.

Severity

The severity of harm to data-protection rights depends materially on the nature of the personal data involved, the nature of the processing, and the scope of any incident affecting that processing. The European Data Protection Board's Guidelines 01/2021 on Examples regarding Personal Data Breach (14 December 2021)[70] provides a spectrum of examples, ranging from incidents giving rise to no significant effect on data subjects (for example, a ransomware attack with proper backup and without exfiltration) to incidents giving rise to high risk to data subjects (for example, highly confidential personal data sent in error).

[70] European Data Protection Board, *'Guidelines 01/2021 on Examples regarding Personal Data Breach'*, 14 December 2021, url.

[redacted]

The processing carried out in connection with Snapchat's in-scope services does not, in the ordinary course, fall within the three high-risk processing examples expressly identified in Article 35(3) GDPR (systematic and extensive evaluation based on automated processing producing legal or similarly significant effects; processing on a large scale of special categories of data or data relating to criminal convictions; systematic monitoring of a publicly accessible area on a large scale).

Notwithstanding these classifications, the volume and continuity of processing means that, were the safeguards regime to fail materially - whether through technical compromise, design error, or operational lapse - the resulting impact on data subjects could be severe.

Recognising both the spectrum of possible harm and the structural exposure inherent in the processing footprint, the severity of negative effects on the Right to the Protection of Personal Data is assessed as **severe harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems process personal data ([redacted] to determine content selection.[redacted] The recommender systems do not, in the ordinary course, knowingly recommend content that violates data-protection rights. Further detail in Section 5.6.
Content moderation systems	Content moderation systems process personal data ([redacted] to enforce the substantive Community Guidelines. [redacted]. The substantive prohibitions on impersonation and on the unauthorised disclosure of personal information (doxxing) are addressed in Sections 4.1.10 and 4.1.7 respectively. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Snap Privacy Policy, the Privacy Center, the Community Guidelines (including the prohibitions on doxxing and impersonation), and the relevant product-specific terms (Spotlight Terms, Lens Studio Terms) together describe the data-protection commitments and the substantive prohibitions applicable to user-generated content engaging the rights of others. Further detail in Section 5.2 and Section 5.5.

[redacted]

Systems for selecting and presenting advertisements	Advertising systems process personal data for ad selection and delivery. [redacted]. Further detail in Section 5.7.
Data-related practices	Privacy by Design is the foundational approach for product and feature development at Snap. Each product and feature is subject to PASS Review (Privacy Assessment System), which is configured to ensure that products do not misuse user data. Data Protection Impact Assessments (DPIAs) are maintained for high-risk processing activities. [redacted] Lenses are designed to use object detection rather than facial identification - Lenses can determine whether a face is present, but do not identify specific faces. Snap's Data Protection Officer function operates as an independent point of contact for users, supervisory authorities and internal stakeholders.
Intentional manipulation (Article 34(2))	[redacted]

[redacted]

Amplification and rapid, wide dissemination	Amplification of data-protection incidents is constrained by the platform architecture (data minimisation principles applied in product design; short default retention for public stories (unless saved to Public Profile); the absence of an open recommender feed for unvetted user-generated content; restrictions on data export; the bidirectional friending requirement that limits the audience for user-disclosed personal information). Where a personal data breach affecting users in the European Union occurs, Snap's incident-response framework provides for notification to the competent supervisory authority and, where applicable, affected data subjects in accordance with Articles 33 and 34 GDPR.
Regional and linguistic aspects	Snap operates as a controller subject to the GDPR in respect of the processing of personal data of EU/EEA users. Snap engages with EU/EEA supervisory authorities, as appropriate, through established GDPR cooperation mechanisms. The Snap Privacy Policy, the Privacy Center, and data subject rights handling pathways are available in multiple EU languages.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, the risk to the Right to the Protection of Personal Data is assessed at **Level 1** overall potential risk prioritization. This output corresponds to the matrix intersection of severe harm industry-wide with highest likelihood. As noted in the Likelihood discussion above, the highest-likelihood determination is driven by structural exposure inherent in the processing footprint rather than by the empirical incidence of complaints. There is no change in this determination from the 2025 Report.



Snap's Mitigations

[redacted]

The mitigation framework for the Right to the Protection of Personal Data is built on Snap's foundational commitment to Privacy by Design and operationalised through a multi-layer architecture spanning product design, governance review, ongoing monitoring, and external transparency. The Article 35(1)(a)–(k) DSA mitigation mapping is recorded below; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. Privacy is one of Snap's foundational platform-governance values, reflected in the architectural design of Snapchat. Specific design features supporting data protection include: data minimisation principles applied in feature design; short default retention for public stories (unless saved to Public Profile); [redacted] the bidirectional friending requirement; private friend lists by default; restrictions on data export; the use of object detection rather than facial identification in Lenses. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Snap Privacy Policy describes the categories of personal data processed, the purposes of processing, the legal bases relied upon, the retention periods, the sharing arrangements with third parties, and the rights available to data subjects. The Community Guidelines prohibit user conduct that engages the data-protection rights of others (in particular, doxxing and impersonation, addressed in Sections 4.1.7 and 4.1.10 respectively). Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Moderation systems process personal data under PASS Review and DPIA scrutiny. The Article 16 DSA notice-and-action mechanism operates as the principal external pathway for receipt of notices alleging illegal content (including content engaging the data-protection rights of others). Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Recommender systems are configured under PASS Review and DPIA scrutiny. Algorithmic systems do not categorise users on the basis of special category data (Article 9 GDPR) for the purpose of advertising targeting, content recommendation or moderation decision-making. Further detail in Section 5.6.

[redacted]

Advertising Systems - Art. 35(1)(e)	Applies. Advertising systems are configured under PASS Review and DPIA scrutiny. Specific data-protection mitigations include: the prohibition on advertising targeting based on minor users' activity data; [redacted] the discontinuation of EU political advertising as of October 2025. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. The volume and nature of privacy operations requests, the volume and nature of queries to the Data Protection Officer flagged as relating to GDPR or Member State privacy laws, the related invasion-of-privacy PVP indicator, and the operational outputs of PASS Review and DPIA processes are monitored on a continuing basis. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Limited application. Article 22 DSA trusted-flagger mechanisms are not principally directed at data-protection complaints, which are handled through the GDPR-aligned data subject rights pathway and through engagement with Member State supervisory authorities. Further detail in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Limited application. No specific code of conduct under Article 45 DSA is currently in operation in respect of data protection on the in-scope services. Snap cooperates with other providers and privacy stakeholders through relevant industry groups, including [redacted] and maintains established protocols for responding to privacy and security incidents, including its Security Incident Response Policy. Snap also engages with EU/EEA data-protection regulators and supervisory authorities through established GDPR cooperation mechanisms.
Transparency - Art. 35(1)(i)	Applies. The Snap Privacy Policy, the Privacy Center, the Privacy Center Help Center articles, the Safety Center, the Just-In-Time (JIT) notices presented during onboarding for features involving data processing (including generative-AI features and My Selfie), and the Transparency Reports together provide information to users on data-protection commitments, on user controls, and on the resources available. Further detail in Section 5.3.

[redacted]

Protection of Minors - Art. 35(1)(j)	Applies. Mitigations specific to minors include: (i) the contact-restriction defaults (Section 5.1.3); (ii) the prohibition on advertising targeting based on minor users' activity data; (iii) the privacy-by-default settings on Snap Map for minor accounts; (iv) Family Center oversight; (v) [redacted]. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. [redacted] The 48-hour input-image deletion for high-latency Lens generation is a specific data-minimisation measure responsive to the data-protection considerations engaged by generative-AI features. Watermarking and contextual identification of generative-AI content support both content-authenticity and data-protection objectives. Further detail in Section 5.9.

PASS Review and the integrated DDIA framework

Two integrated governance instruments are central to Snap's data-protection mitigation framework and warrant specific reference:

- **PASS Review (Privacy Assessment System).** Each product and feature is subject to PASS Review at appropriate stages of development. [redacted]
- **DDIA framework (Section 6.3).** The Digital and Data Impact Assessment framework integrates [redacted]

[redacted]

Together, the PASS Review and the DDIA framework operationalise Article 25 GDPR (Data Protection by Design and by Default) and Article 35 GDPR (DPIA) requirements as integrated components of Snap's product-development workflow.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place in respect of the Right to the Protection of Personal Data on Snapchat's in-scope services are reasonable, proportionate and effective.

The Level 1 overall risk prioritisation reflects the structural exposure inherent in the volume of processing carried out in connection with Snapchat's in-scope services. The mitigation framework responds to that structural exposure through: (i) Privacy by Design as a foundational platform-governance value reflected in architectural decisions (short default retention for public stories (unless saved to Public Profile), [redacted] the use of object detection rather than facial identification in Lenses); (ii) the integrated PASS Review and DDIA framework operationalising Articles 25 and 35 GDPR through pre-deployment governance review; (iii) the Snap Privacy Policy and Privacy Center providing transparent user-facing information; (iv) the Data Protection Officer function as an independent point of contact for users and supervisory authorities; (v) data subject rights handling pathways aligned with GDPR Articles 15–22; (vi) the Article 33/34 GDPR-aligned incident response framework; and (vii) ongoing engagement with EU/EEA data-protection supervisory authorities through the established GDPR cooperation mechanisms.

The empirical observations recorded above (low absolute volume of privacy-related queries against the population of approximately 96.5 million EU monthly active recipients) support the determination that the safeguards regime has, in operation, kept the user-facing incidence of data-protection harm low. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to the data-protection implications of generative-AI features and to the operational outputs of EU AI Act compliance work where applicable. There is no change in the overall conclusion from the 2025 Report.

4.2.5 Right to Non-Discrimination and Freedom of Religion

This Section 4.2.5 sets out Snap's assessment of the actual or foreseeable risk of negative effects on the Right to Non-Discrimination and the Right to Freedom of Religion arising from the

operation of Snapchat's in-scope services. Both rights are recognised under the EU Charter (Articles 10, 21 and 22) and the European Convention on Human Rights (Articles 9 and 14). The substantive risk pathways for this risk are addressed in detail in two underlying Category 1 Sections, Section 4.1.2 (Hate Speech) and Section 4.1.4 (Terrorist Content), and in the algorithmic-bias dimensions of Snap's recommender, moderation and advertising systems.

Online platforms can be used to spread content containing or promoting discrimination on grounds of sex, racial or ethnic origin, nationality, religion or belief, disability, age, sexual orientation or other identity markers; can host algorithmic systems that, in the absence of mitigation, could produce or amplify discriminatory outcomes; can deliver advertising in ways that produce discriminatory exposure or exclusion; and can host reporting mechanisms susceptible to bad-faith use against accounts associated with marginalised communities. This Section assesses the residual risk to the Right to Non-Discrimination and the Right to Freedom of Religion given the operational mitigation framework in place across the in-scope services.

minor-account default settings (Section 5.1.3), the friending-protection refinements (Section 5.4.4), and the profile-level moderation expansion (Section 5.4.4) reduce the contact and exposure pathways for discriminatory conduct directed at minors and at users from protected groups. The Article 16 and Article 23 reporting-pathway and misuse-suspension changes (Section 5.4.6) support the reactive-reporting dimension across categories including hate-related categories.

Likelihood

As recorded in the 2025 Report, in the absence of safeguards the dissemination of Hate Speech and content that demeans, defames or promotes discrimination or violence on the basis of protected-class status would be a high-likelihood risk on any platform that hosts user-generated content. The likelihood determination for this risk is therefore driven by the operational performance of the safeguards regime in place, rather than by the absence of underlying threat actors.

The likelihood determination for this risk draws on three principal evidence bases:

- **Underlying Category 1 likelihood determinations.** The likelihood of the Category 1 risks that bear directly on this fundamental right are recorded as follows:

Underlying Category 1 risk	2026 likelihood
Section 4.1.2 - Dissemination of Illegal Hate Speech	Lowest likelihood

Section 4.1.4 - Dissemination of Terrorist Content	Lowest likelihood

No separate PVP or VVR is presented for the Right to Non-Discrimination and the Right to Freedom of Religion because this is a rights-based aggregation of content-related and algorithmic-bias risks rather than a standalone policy or enforcement category. Relevant PVP and VVR evidence is presented in the corresponding Category 1 Sections, principally Sections 4.1.2 and 4.1.4.

- **Algorithmic-bias-related feedback and complaints.** Snap maintains privacy, support and complaint-handling channels through which users can raise concerns regarding potentially discriminatory outcomes associated with algorithmic systems. These channels provide signals that may support the identification of recurring or systemic bias-related issues. No quantitative conclusion is drawn from bias-specific complaint volumes for this assessment period.
- **EU Code of Conduct monitoring outcomes.** In the first monitoring exercise under the revised EU Code of Conduct on Countering Illegal Hate Speech Online, conducted from November to December 2025 with results published in April 2026, Snapchat received no relevant notifications from the participating monitoring organisations. The revised exercise applied an updated methodology and is not directly comparable with the earlier monitoring rounds. Further detail is provided in Sections 4.1.2 and 5.12.

On the basis of the underlying Category 1 likelihood determinations, lowest likelihood for both Hate Speech and Terrorist Content, the revised-Code monitoring outcome and the safeguards applying to algorithmic systems, the likelihood of negative effects on the Right to Non-Discrimination and the Right to Freedom of Religion on Snapchat's in-scope services is assessed at the **lowest likelihood** category. No quantitative conclusion is drawn from algorithmic-bias-specific complaint volumes for this assessment period. There is no change in this determination from the 2025 Report.

Severity

Discrimination and persecution on grounds of protected characteristics pose a serious threat to the rights of EU citizens, many of whom belong to communities that have historically faced discrimination and marginalisation. External evidence supports treating discriminatory algorithmic outcomes as a serious industry-wide concern. The EU Agency for Fundamental Rights has documented how bias can arise in algorithmic systems and has emphasised the importance of testing and ongoing monitoring to identify and mitigate potentially discriminatory effects

- The Netherlands Institute for Human Rights found that the Breeze dating app's matching algorithm discriminated against non-white users, resulting in a formal order to address the bias.[71]
- The European Commission has demanded internal documentation from X (formerly Twitter) on how its recommendation systems may be amplifying far-right content, underscoring concerns about algorithmic moderation potentially skewing democratic discourse.[72]

These examples reflect the scrutiny that algorithmic bias has attracted within both legal frameworks and regulatory action, particularly under the evolving mandates of the EU AI Act and the GDPR.

The severity of harm associated with discriminatory content and discriminatory algorithmic outcomes is assessed by reference to the underlying Category 1 severity classifications: Hate Speech is classified as significant harm industry-wide (Section 4.1.2) and Terrorist Content as serious harm industry-wide (Section 4.1.4). Taking the higher of these as the severity reference point for this risk, and recognising the broader societal-level harm that algorithmic bias and structural discrimination can produce, the severity of negative effects on the Right to Non-Discrimination and the Right to Freedom of Religion is assessed as **serious harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to knowingly recommend hateful or discriminatory content; no corresponding interest categories exist. Algorithmic systems do not categorise users on the basis of special category data (Article 9 GDPR) for the purpose of advertising targeting, content recommendation or moderation decision-making. Recommender system design is subject to PASS Review and DPIA scrutiny (Section 4.2.4), with bias testing forming part of the development pipeline. Further detail in Section 5.6.

[71] Tim de Jonge, Frederik Zuiderveen Borgesius, 'Mitigating Digital Discrimination in Dating Apps -- The Dutch Breeze case', url, (2024).
[72] Financial Times, 'Brussels orders X to hand over documents on algorithm', url, (2025).

[redacted]

Content moderation systems	The proactive and reactive moderation framework operating against Hate Speech (Section 4.1.2) and Terrorist Content (Section 4.1.4) is the principal operational pathway for content directly engaging the Right to Non-Discrimination. [redacted] Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Hateful Content, Terrorism and Violent Extremism prohibit Hate Speech and the use of hate symbols, and the Community Guidelines apply uniformly across in-scope services and equally to all users. The substantive prohibitions cover content that demeans, defames or promotes discrimination on protected grounds (race, colour, caste, ethnicity, national origin, religion, sexual orientation, gender, gender identity, disability, veteran status, immigration status, socio-economic status, age, weight or pregnancy status). Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Specific anti-discrimination measures apply to advertising. [redacted] Further detail in Section 5.7.
Data-related practices	Privacy and safety by design practices (Section 4.2.4) apply to product and feature changes that may bear on this risk, including in respect of the design of algorithmic systems where bias-related considerations engage. The PASS Review and DDIA framework include consideration of fundamental-rights impacts (including non-discrimination impacts) for high-risk processing activities. Where applicable, EU AI Act compliance work addresses the bias-related dimensions of algorithmic systems within the AI Act risk-classification framework.

[redacted]

Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of discriminatory content on in-scope services is structurally constrained by the platform architecture (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation) and by the explicit configuration of recommender systems not to surface hate-related content. The Content Guidelines for Recommendation Eligibility set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution.
Regional and linguistic aspects	ALD banned-word lists for hate-related content are maintained across multiple languages and categorised by context and region. Moderation guidance is calibrated to Member State legal contexts, including in respect of categories of content that are illegal in specific Member States (for example, specific symbols, denial of historical atrocities, and specific protected-class slurs). Member State engagement on anti-discrimination measures is maintained, including (in 2024) Snap's participation in a roundtable hosted by the Austrian Minister for EU and Constitutional Affairs and the President of the Jewish Religious Community of Austria, addressing measures to curb the spread of antisemitic hate content online.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, the risk to the Right to Non-Discrimination and the Right to Freedom of Religion is assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of serious harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

	Lowest likelihood	Medium likelihood	Highest likelihood

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. The mitigation framework operates principally through the underlying Category 1 mitigation set (Hate Speech, Section 4.1.2; Terrorist Content, Section 4.1.4), supplemented by the algorithmic-bias-specific measures in advertising systems and recommender systems. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. The platform-architecture factors that constrain amplification across all categories (recommended content in public feed is moderated; short default retention for public stories (unless saved to Public Profile); no broad reshare functionality of Stories; pre-distribution moderation for content reaching wide audiences) operate to limit the structural opportunity for dissemination of discriminatory content. Programmes to foster a diverse content community (including content-creator support programmes) operate alongside the substantive prohibitions to support a positive user experience for users from a range of backgrounds. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Hateful Content, Terrorism and Violent Extremism (and the related substantive prohibitions across other Community Guidelines categories) prohibit discriminatory content as described in the DSA Risk Factors above. The Community Guidelines apply uniformly across in-scope services and equally to all users. Further detail in Section 5.2 and Section 5.5.

[redacted]

Moderation - Art. 35(1)(c)	Applies. Proactive moderation of hate-related content is recorded in detail in Section 4.1.2 and Section 5.4. [redacted] Reactive moderation is available through the in-app reporting flow with a dedicated Hate Speech reporting reason. The Article 23 DSA misuse-suspension framework (Section 5.4.6) addresses bad-faith reporting patterns including mass-reporting of marginalised-community accounts. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Hate-related or discriminatory content is not eligible for recommendation, and no corresponding recommendation-interest categories exist. Bias and non-discrimination considerations form part of the development and review of relevant algorithmic systems through PASS Review and, where applicable, DPIA scrutiny. Applicable EU AI Act compliance processes provide an additional framework for assessing bias-related risks in relevant systems. Further detail is provided in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Specific anti-discrimination measures apply to advertising: [redacted] and the Advertising Policies prohibit advertising that contains discriminatory content. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Hate Speech (Section 4.1.2) and Terrorist Content (Section 4.1.4) provides the principal data inputs for monitoring this risk. Privacy, support and complaint-handling channels are also monitored for recurring or systemic concerns relating to potentially discriminatory outcomes associated with algorithmic systems. No quantitative conclusion is drawn from algorithmic-bias-specific complaint volumes for this assessment period. Further detail is provided in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Cooperation with designated trusted flaggers under Article 22 DSA in relation to hate-related content is described in Section 4.1.2, including engagement with Licra in France, the German Department for Internet Services & Social Media and Someturva in Finland. Notices submitted by designated trusted flaggers are processed through the applicable priority-notice procedures. Further detail is provided in Section 5.10.

[redacted]

Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap is a signatory of the EU Code of Conduct on Countering Illegal Hate Speech Online (revised 20 January 2025; recorded in Section 4.1.2). Snap engages with civil-society organisations focused on anti-discrimination and human rights, including [redacted] and other partner organisations. Member State-specific engagement is maintained, including the 2024 Austrian roundtable engagement noted above. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Hateful Content, Terrorism and Violent Extremism category explainer, the Safety Center, and Snap's Transparency Reports provide information on the substantive prohibitions, on enforcement and prevalence, and on the resources available to users. The diverse-content-community programmes (including content-creator support programmes) are recorded in Snap's external communications. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Mitigations specific to minors include: (i) the contact-restriction defaults (Section 5.1.3); (ii) Family Center oversight (Section 5.8); (iii) the prohibition on advertising targeting based on minor users' activity data (Section 4.2.4 and Section 5.7); and (iv) [redacted] Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Partially applies. The Community Guidelines apply without regard to whether discriminatory content is user-generated or AI-generated. Content-authenticity measures (watermarking and contextual identification of generative-AI content) operate alongside the substantive prohibitions and are described in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the protection of the Right to Non-Discrimination and the Right to Freedom of Religion on Snapchat's in-scope services are reasonable, proportionate and effective.

The mitigation framework operates principally through the underlying Category 1 mitigation set for Hate Speech (Section 4.1.2) and Terrorist Content (Section 4.1.4), and through the algorithmic-bias-specific measures in advertising and recommender systems. The combination of:

[redacted]

the substantive Community Guidelines prohibitions on hate-related content; the proactive moderation suite [redacted]; the prohibition on advertising-targeting based on minor users' activity data; the EU Code of Conduct on Countering Illegal Hate Speech Online signatory status; cooperation with designated trusted flaggers under Article 22 DSA in relation to hate-related content; and the engagement with civil-society organisations and Member State authorities (including the 2024 Austrian roundtable engagement) supports the conclusion that the mitigation framework operates effectively against this risk.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to the bias-related dimensions of algorithmic systems as the EU AI Act compliance framework matures, and to the operational outputs of any 2026 engagement with anti-discrimination civil-society partners or Member State authorities.

4.2.6 Children's Rights

This Section 4.2.6 sets out Snap's assessment of the actual or foreseeable risk of negative effects on Children's Rights arising from the operation of Snapchat's in-scope services.

Scope of this risk

The "rights of the child" under Article 24 of the EU Charter comprises two principal elements relevant to Snapchat's in-scope services:

- **Element 1 - Right to protection and care.** Children have the right to such protection and care as is necessary for their well-being. The substantive risks engaging this element are addressed comprehensively in **Section 4.4.3 (Negative Effects on Minors)** and **Section 4.4.4 (Negative Effects on Physical and Mental Well-Being)**.
- **Element 2 - Right to express views.** Children have the right to express their views freely and to have those views taken into consideration on matters which concern them, in accordance with their age and maturity. This Section 4.2.6 focuses on this second element.

Accordingly, this section assesses the risk to children's right to participation and expression on Snapchat's in-scope services. The substantive protection-and-care dimension is addressed in Sections 4.4.3 and 4.4.4, with the associated minor-protection framework recorded in Section 5.8; this section provides the fundamental-rights assessment of children's expression and participation.

EU minor user population

[redacted]

As recorded in the Snapchat Community sub-section of Section 2, the average monthly active recipient population in the European Union included [redacted] users in the 13–17 age cohort, representing approximately [redacted] of the total EU monthly active recipient population on Snapchat. [redacted]

Likelihood

Consistent with the scope of this risk, the likelihood determination below concerns the participation-and-expression dimension of Children's Rights. The protection-and-care dimension is assessed separately in Sections 4.4.3 (Negative Effects on Minors) and 4.4.4 (Negative Effects on Physical and Mental Well-Being), and the likelihood determinations for those risks are not repeated here. For the participation-and-expression dimension, the relevant proxy is the underlying Section 4.2.2 (Right to Freedom of Expression and Assembly) determination, which assesses the likelihood of negative effects on freedom of expression at the lowest likelihood category for the Snapchat community generally.

Risk category	Likelihood - Snapchat community generally	Likelihood - children specifically
Right to Freedom of Expression (Section 4.2.2)	Lowest likelihood	Lowest likelihood

The likelihood conclusion is supported by the following:(i) the underlying minor-specific Violative View Rate figures across the substantive harm categories recorded in Sections 4.1.2, 4.1.4, 4.1.6, 4.1.7, 4.1.9 and 4.1.10, which consistently show very low minor-specific exposure rates; (ii) the design and procedural safeguards described in Sections 4.2.2, 4.4.3 and 5.8; and (iii) the contact-restriction defaults and Family Center framework supporting safe digital participation by minor users. No quantitative conclusion is drawn from minor-specific complaint volumes for this assessment period.

External evidence

The contemporary academic and regulatory record documents an evolving discussion on the appropriate balance between minor-protection measures and children's right to digital participation. Selected sources include:

- **LSE EU Kids Online research**. LSE EU Kids Online research (February 2025) on the children's-rights implications of EU social-media restrictions, identifying that overly broad age-based restrictions can themselves engage children's rights to expression, information and participation;

- **Eric Goldman**. Eric Goldman, “Segregate-and-Suppress: The Constitutional and Other Implications of Age-Based Online Speech Restrictions” (SSRN), identifying that overly restrictive minor-targeted speech regulations may infringe minors’ fundamental rights to expression and access to information;
- **ThinkYoung research**. ThinkYoung research (February 2025) on European youth perspectives on online platforms, including direct evidence of how teenagers themselves perceive platform participation and its role in their development and social connection;
- **University of Amsterdam**. University of Amsterdam research (2024) on Snapchat-specific findings, identifying positive associations between platform use and adolescent social connection and well-being in the studied populations.

These sources support the analytical position that Children’s Rights encompass both protection-and-care and participation-and-expression. The former is assessed substantively in Sections 4.4.3 and 4.4.4; this section assesses the latter. They also indicate that measures intended to protect minors may themselves affect children’s access to information, expression and participation if they are broader or more restrictive than necessary.

Likelihood conclusion

On the basis of the underlying Section 4.2.2 likelihood determination, the minor-specific exposure data (VVR) and the safeguards supporting safe digital participation, the likelihood of negative effects on the participation-and-expression dimension of Children’s Rights on Snapchat’s in-scope services is assessed at the **lowest likelihood category**.

Severity

The severity of harm associated with negative effects on Children’s Rights warrants elevated assessment relative to the same risks affecting adult users, on two analytical bases:

- **Vulnerability of the population.** Children and teens (13–17) are recognised as a vulnerable population by virtue of their developmental stage, evolving cognitive and emotional capacities, and limited capacity to assess and respond to certain online risks. Where harm materialises, the impact on a minor can be more severe than the equivalent harm to an adult, and can have lasting effects on development, education, and digital confidence (as recorded across the substantive Category 1 and Category 4 sections).
- **Importance of the right.** The right of children to express views and to participate in digital life is itself a fundamental right under Article 24 of the EU Charter, Articles 12 and 13 UNCRC, and General Comment No. 25 of the UN Committee on the Rights of the Child (2021). It is essential that children and teens are able to access online platforms and to participate in lawful online debate and dialogue - to learn, to have their views heard, to develop their values and identities - regardless of their ability to pay. Restrictions that infringe this right (whether through inappropriate content removal, inappropriate restriction of contact, or other measures) themselves constitute a material harm.

[redacted]

On these analytical bases, the severity of negative effects on Children's Rights is assessed as **severe harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services. The substantive treatment of the risk factors as they bear on the protection-and-care dimension of Children's Rights is recorded in Section 4.4.3 (Negative Effects on Minors); the treatment below focuses on the participation-and-expression dimension.

Design of recommender systems and other algorithmic systems	Recommendation eligibility and ranking for public content are subject to additional standards that account for the fact that users as young as 13 may view the content. These standards are intended to reduce minors' exposure to harmful or inappropriate content without categorically restricting their access to lawful expression and information. Recommendation operates only on defined public-content surfaces and only in respect of content satisfying the applicable recommendation-eligibility requirements. Further detail is provided in Sections 4.4.3 and 5.6.
Content moderation systems	Substantive moderation framework recorded in Section 4.4.3 (Minors) and Section 5.4. The reactive reporting pathway includes the dedicated "It involves a child" reporting reason, supporting the child's ability to report content engaging their rights.
Applicable terms and conditions and enforcement	Snap's Terms of Service and Community Guidelines apply uniformly across in-scope services and equally to all users including minors. Specific content prohibitions calibrated to the child-population baseline (in particular, the assumption in the Content Guidelines for Recommendation Eligibility that users as young as 13 may be viewing content) are recorded in Section 4.4.3 and Section 5.2. Enforcement of that child-population baseline reduces the risk that recommendation-eligible public content exposes minors to age-inappropriate material, while preserving minors' access to lawful content, information and expression. Teen-friendly Terms language and support pages are maintained to support comprehension by users in the 13–17 age cohort.

[redacted]

Systems for selecting and presenting advertisements	Specific minor-protection measures applicable to advertising are recorded in Section 4.4.3 and Section 5.7, and include: the prohibition on advertising targeting based on minor users' activity data; [redacted] the discontinuation of EU political advertising as of October 2025; [redacted]
Data-related practices	PASS Review and DDIA framework apply enhanced minor-protection scrutiny to product and feature changes that affect minor users (recorded in Section 4.2.4 and Section 6.3). The integration of the Age Appropriate Design Code (AADC) Risk Assessment within the DDIA framework operationalises minor-specific data-protection requirements as part of product-development workflow.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	The platform-architecture factors that constrain amplification across all categories operate to limit the structural opportunity for content infringing children's rights. The bidirectional friending requirement and the contact-restriction defaults for minor accounts particularly constrain unsolicited contact-based amplification pathways.
Regional and linguistic aspects	Snap follows age-appropriate design frameworks in jurisdictions where these are established, including the United Kingdom, France and California, and applies the underlying principles globally as a baseline. EU/EEA-specific work includes assessment against the European Commission's Article 28 Guidelines and engagement with relevant child-protection authorities through established channels. Further detail is provided in Sections 5.8 and 6.

Overall Potential Risk Prioritization

The risk-matrix output for this risk is severe harm industry-wide × lowest likelihood = **Level 3** by matrix default. However, applying the case-by-case deviation framework recorded in Section 3.5,

[redacted]

this risk is assessed at **Level 1** overall potential risk prioritization. The deviation up from Level 3 to Level 1 is supported by the two analytical bases recorded in the Severity discussion above:

- the inherent vulnerability of the minor-user population, recognised across the EU Charter, UNCRC, GDPR (recital 38, recital 75; Article 8), and substantive case law of the CJEU; and
- the foundational importance of children's rights to expression and digital participation, recognised across the same instruments and given specific operational substance by General Comment No. 25 of the UN Committee on the Rights of the Child (2021).

This deviation pattern is consistent with the analytical treatment of Section 4.1.1 (CSEA Content) and Section 4.1.8 (Self-Harm), each of which similarly deviates upward from the matrix default to Level 1 on a case-by-case basis under Section 3.5. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. The substantive minor-protection framework is recorded in detail in Section 5.8 (Protection of Minors); the mitigation table provides the participation-and-expression view of the framework, with cross-references to the substantive treatment in Section 5.8 and Section 4.4.3.

Snapchat Design and Function - Art. 35(1)(a)	Applies. The platform-architecture factors that support children's safe digital expression include: bidirectional friending requirement; contact-restriction defaults for minor accounts; short default retention for public stories (unless saved to Public Profile); the absence of an open recommender feed for unvetted user-generated content. The minor-account default settings (Section 5.1.3) and [redacted] strengthen the design framework. Substantive detail in Section 4.4.3 and Section 5.1.

[redacted]

Terms and Enforcement - Art. 35(1)(b)	Applies. Snap's Community Guidelines and Terms of Service apply uniformly to minors. Teen-friendly Terms language and support pages are maintained to support comprehension by users in the 13–17 cohort. The substantive prohibitions (recorded across Sections 4.1.1–4.1.12) operate to prevent content that would infringe children's rights or restrict their safe digital participation. Substantive detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. The proactive and reactive moderation framework is recorded in Section 4.4.3 and Section 5.4. The reactive reporting pathway includes the dedicated "It involves a child" reporting reason. The expansion of profile-level moderation (Section 5.4.4) supports detection of patterns engaging children's rights at the username, display-name and profile level.
Algorithmic Systems - Art. 35(1)(d)	Applies. Recommender systems are configured to apply the Content Guidelines for Recommendation Eligibility, which set additional standards for content eligible for algorithmic distribution - with the operational assumption (in particular for Discover) that users as young as 13 may be viewing content. Substantive detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Specific minor-protection advertising measures are recorded in Section 4.4.3 and Section 5.7, including: the prohibition on advertising targeting based on minor users' activity data; [redacted] the discontinuation of EU political advertising as of October 2025; [redacted] Substantive detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. The Section 6 ongoing risk-detection processes - including consultation with the Safety Advisory Board (which includes child-safety experts), engagement with civil-society organisations focused on child protection, the PASS Review and DDIA framework, Prevalence Testing and, separately, minor-specific Violative View Rate tracking - monitor risks bearing on this category. Substantive detail in Section 6.

[redacted]

Trusted Flaggers - Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA may concern substantive harms affecting minors and are processed through the applicable priority-notice procedures. Snap also cooperates with specialist child-safety organisations, including NCMEC and members of the INHOPE network, through the separate operational frameworks described in Sections 4.1.1 and 5.10. These mechanisms support the broader minor-protection framework and, in turn, children's safe digital participation. Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap engages with the [redacted] with civil-society organisations focused on children's rights, and with Member State child-protection frameworks. [redacted] Substantive detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. Information for minors and their parents/guardians is provided through: the Teen-friendly Community Guidelines and Terms summary; the Family Center support pages; the parents site; the Safety Center; the Privacy Center; and Just-In-Time (JIT) notices presented during onboarding for features involving data processing. The Transparency Reports include minor-specific data where relevant (in particular, minor-user Violative View Rate figures across substantive harm categories). Substantive detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. The full minor-protection framework is recorded in Section 5.8 (Protection of Minors), which addresses: app-store level safeguards; device-level safeguards; platform-level safeguards; product-level safeguards (Lenses, Spotlight, Discover, Public Profiles, Snap Map, advertising); reporting and blocking; private messaging and friending controls; Family Center / parent tools; and the Article 28 Guidelines mapping. The Section 5.8 Article 28 Guidelines mapping is the principal substantive reference point for assessment of measures against the Commission's Guidance.
Content Authenticity - Art. 35(1)(k)	Applies. Generative-AI features made available to minor users are subject to the additional content-authenticity safeguards recorded in Section 5.9, including watermarking, retry-logic for unsafe outputs, and JIT notices warning users about the AI-generated character of outputs. Substantive detail in Section 5.9.

Article 28 Guidelines (Commission Guidance, 2025)

Following the publication by the European Commission of its Guidance on Article 28 DSA in 2025, Snap has assessed Snapchat's in-scope services against the Guidance's principal recommendations. The detailed assessment is recorded in Section 5.8 (Protection of Minors), where Snap's measures are mapped against the eight principal sections of the Guidance (General Principles; Risk Review; Age Assurance; Defaults and Account Settings; Interfaces, Tools and AI; Recommender Systems and Content Moderation; Commercial Practices; Reporting Support and Guardian Tools). The detailed Section 5.8 mapping is the principal substantive reference point for the Article 28 Guidelines work.

The detailed assessment in Section 5.8 records how Snap's existing measures correspond to the principal recommendations in the Article 28 Guidelines, including areas in which existing safeguards are considered proportionate to the nature of Snapchat's in-scope services and recommendations that remain under continued consideration. Section 5.8 is the controlling reference for the status of that assessment. Snap continues to engage with the European Commission regarding implementation of the Guidelines.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors and the mitigations identified above, Snap has concluded that the mitigations addressing the participation-and-expression dimension of Children's Rights on Snapchat's in-scope services are reasonable, proportionate and effective. The protection-and-care dimension is assessed principally in Sections 4.4.3 and 5.8.

The framework operates principally through the substantive minor-protection measures recorded in Section 4.4.3 and Section 5.8, supplemented by the architectural and policy commitments that support children's safe digital participation: the bidirectional friending requirement; the contact-restriction defaults for minor accounts; the prohibition on advertising targeting based on minor users' activity data; the Content Guidelines for Recommendation Eligibility calibrated to a child-population baseline; the dedicated reporting pathways including the "It involves a child" reporting reason; the Teen-friendly Terms language and support pages; the Family Center oversight framework; the integration of the Age Appropriate Design Code Risk Assessment within the DDIA framework; and the engagement with the Commission and Member State authorities on the Article 28 Guidelines.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to: the operational outputs of the minor-account default and onboarding refinements; the operational implementation of the Article 28 Guidelines as the Commission's

expectations evolve; the academic and civil-society research base on the appropriate balance between protection and participation in respect of children's digital lives; and the continued integration of the Age Appropriate Design Code framework into Snap's DDIA processes.

Snap has continued to actively support the Commission's work and the Member State frameworks on minor protection, including through the Article 28 Guidelines engagement, the engagement with EU/EEA child-protection authorities, and the engagement with civil-society organisations focused on children's digital rights. This engagement will continue through subsequent assessment periods.

4.2.7 Right to Consumer Protection

This Section 4.2.7 sets out Snap's assessment of the actual or foreseeable risk of negative effects on the Right to Consumer Protection arising from the operation of Snapchat's in-scope services. The Right to Consumer Protection is recognised under Article 38 of the EU Charter and Article 169 TFEU, and given operational substance by a body of EU consumer-protection legislation including the Unfair Commercial Practices Directive, the Consumer Rights Directive, and the Audiovisual Media Services Directive (which addresses advertising transparency).

Online platforms can pose risks to consumer-protection rights through a number of pathways: the dissemination of deceptive content (whether user-generated or commercial); the delivery of misleading or fraudulent advertising; manipulative design patterns that materially distort users' free and informed decision-making; processing or sharing of user data in ways that produce disproportionate personalised commercial targeting; and inadequate transparency in respect of commercial communications. This section assesses the risk to the Right to Consumer Protection given the operational mitigation framework in place across the in-scope services.

Likelihood

Internal data
The likelihood determination for this risk derives from three principal evidence bases:

Underlying Category 1 and Category 2 risk determinations
The substantive risk pathways for this risk are addressed in the following underlying sections, with the 2026 likelihood and risk-prioritisation determinations recorded:

Underlying risk	2026 likelihood
Section 4.1.3 - Sale of Prohibited Products (other goods)	Lowest

[redacted]

Section 4.1.10 - Harmful False Information	Lowest
Section 4.1.11 - Fraud and Spam	Lowest
Section 4.2.3 - Right to Private Life	Lowest
Section 4.2.4 - Right to Data Protection	Highest

The relevant PVP and VVR evidence is presented in the underlying Category 1 sections; this section supplements that evidence with consumer-protection-specific advertising review, reporting and removal data.

Of these, the Right to Data Protection (Section 4.2.4) carries highest likelihood by virtue of the structural-exposure rationale recorded in that Section. The available user-facing data nevertheless indicates a comparatively low absolute volume of privacy-related requests. Snap received approximately [redacted] EU Privacy Operations requests during July 2023–June 2024, [redacted] during July 2024–June 2025, and [redacted] during July 2025–June 2026, against an average population of approximately 96.5 million EU monthly active recipients. The latest assessment-period volume therefore remained low in absolute terms and declined relative to each of the preceding two annual periods. These observations do not displace the structural-exposure rationale underlying the highest-likelihood classification in Section 4.2.4, but provide empirical context regarding the observed incidence of user-facing privacy concerns. The other underlying risks are recorded at lowest likelihood

Internal data - advertising-side

Pre-publication review and post-publication reporting data demonstrate the scale of Snap's advertising moderation framework. Between 1 July 2025 and 30 June 2026, [redacted] distinct advertisements completed final review, compared with [redacted] during the preceding 12-month period, representing an increase of approximately [redacted] During the same period, [redacted] advertisements were rejected at final review, compared with [redacted] during the prior comparable period. Fraud- and schemes-related rejection reasons remained a major component of the overall rejection volume.

Separately, Snap's H2 2025 Transparency Report recorded 81,976 advertisements reported following publication, of which 47,266 were removed. These figures reflect the continued operation of reactive reporting and post-publication enforcement alongside the pre-publication review framework.[73]

[73] https://values.snap.com/privacy/transparency-h2-2025.

[redacted]

External evidence

External evidence relevant to this risk includes:

- **European Commission Consumer Conditions Scoreboard.** The European Commission's 2025 Consumer Conditions Scoreboard found that online shoppers were more than 60% more likely to experience purchasing problems than consumers shopping offline. It also found that 45% of consumers had encountered online scams and that 93% of online shoppers had concerns regarding targeted advertising, including the collection of personal data, excessive advertising and personalisation[74].
- **EU consumer-protection enforcement.** Coordinated enforcement by the European Commission and the Consumer Protection Cooperation Network during 2025 identified online consumer-protection concerns including fake discounts, pressure-selling techniques, misleading or incomplete information, deceptive product labels, misleading sustainability claims, and potentially misleading rankings, reviews and ratings. These findings illustrate the range of consumer-protection risks that can arise through online commercial interfaces[75].
- **ESMA and IOSCO engagement on online financial misconduct.** In May 2025, ESMA wrote to Snap and a number of other online-platform providers encouraging proactive measures to prevent the promotion of unauthorised financial services. ESMA identified the increasing use of digital platforms to advertise unlawful financial services as a source of financial loss and reduced trust for retail consumers. The engagement complemented IOSCO's 2025 work on online financial harm and the responsibilities of platform providers.[76]

Likelihood conclusion

On the basis of the underlying Category 1 and Category 2 risk determinations, considered in the context of the specific consumer-protection pathways, the scale and operation of the advertising pre-publication review pipeline, the continued operation of post-publication reporting and enforcement, and the advertising-transparency, targeting and design controls recorded below, the likelihood of negative effects on the Right to Consumer Protection on Snapchat's in-scope services is assessed at the **lowest likelihood category**.

Severity

Negative effects on consumer-protection rights can produce a range of harms, including:

[74] European Commission, Consumer Conditions Scoreboard 2025, 14 March 2025. The Scoreboard recorded that online shoppers were over 60% more likely to encounter purchasing problems, 45% had encountered online scams, and 93% expressed concerns about targeted advertising.
[75] European Commission and Consumer Protection Cooperation Network, coordinated consumer-protection action concerning SHEIN, 26 May 2025.
[76] European Securities and Markets Authority, ESMA urges social media companies to tackle unauthorised financial ads, 28 May 2025; IOSCO, Statement on Combatting Online Harm and the Role of Platform Providers, 21 May 2025.

- direct financial loss to users from fraudulent transactions, deceptive marketing or unsubstantiated commercial claims;
- disproportionate personalised commercial targeting based on extensive user data;
- misleading or false advertisements that cause uninformed purchasing decisions or erode trust in legitimate commercial activity;
- manipulative design patterns that materially distort users' free and informed decision-making (engaging Article 25 DSA);
- inadequate transparency in respect of commercial communications, undermining users' ability to evaluate the source and basis of advertising content (engaging Article 26 DSA);
- particular vulnerability of minor users to commercial targeting and marketing practices.

The European Union has recognised these harms across a body of consumer-protection legislation (the Unfair Commercial Practices Directive, the Consumer Rights Directive, the Audiovisual Media Services Directive in respect of advertising transparency, the Political Advertising Regulation in respect of political advertisements). Member State legislators have similarly developed targeted frameworks (including, by way of example, the Dutch ACM guidelines on transparent and fair online platform economies). Recognising both the financial and broader user-trust dimensions of consumer-protection harm, the severity of negative effects on the Right to Consumer Protection is assessed as **significant harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to knowingly recommend deceptive content, scam content or fraudulent advertising; the platform-architecture factors that constrain amplification (recommended content in public feed is moderated, no broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), pre-distribution moderation) operate to limit the structural opportunity for consumer-protection harm. Further detail in Section 5.6.

[redacted]

Content moderation systems	The proactive and reactive moderation framework operating against Fraud and Spam (Section 4.1.11), Harmful False Information (Section 4.1.10) and Sale of Prohibited Products (Section 4.1.3) is the principal operational pathway for consumer-protection-relevant content. [redacted] The Article 16 DSA notice-and-action mechanism (Section 5.4.6) operates as the principal external pathway for receipt of notices alleging illegal content (including content engaging consumer-protection law). Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Harmful, False or Deceptive Practices and the Advertising Policies prohibit the categories of conduct most directly engaging consumer-protection rights: fraud, scams, deceptive marketing, impersonation, unsubstantiated claims, manipulative practices, and spam. The Commercial Content Policy requires creators and partners to label paid or sponsored content clearly. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Pre-publication review applies to all advertising; [redacted] Specific consumer-protection measures include: the prohibition on advertisements containing unsubstantiated or false claims; the requirement for prominent ad markings supporting Article 26 DSA transparency; [redacted] the discontinuation of EU political advertising as of October 2025 consistent with Regulation (EU) 2024/900; and the prohibition on advertising-targeting based on minor users' activity data. Further detail in Section 5.7.
Data-related practices	PASS Review and DDIA framework (Section 6.3) apply to product and feature changes that bear on consumer-protection-relevant data processing, including in respect of advertising systems and personalisation. The Snap Privacy Policy and the Privacy Center provide user-facing transparency on data use, with opt-out controls for personalisation where required. Further detail in Section 4.2.4.

[redacted]

Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of consumer-protection-relevant content is structurally constrained by the platform architecture (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences; pre-publication review for all advertising). The advertising review pipeline rejects more than [redacted] advertisements per reporting period, with fraud- and schemes-related categories comprising a major component of the overall rejection volume (further detail in Section 4.1.11).
Regional and linguistic aspects	Member State consumer-protection frameworks and enforcement priorities vary across the EU. Snap's principal consumer-protection controls, including the Article 26 DSA advertisement-transparency requirements, operate across Member States. Snap also considers guidance issued by EU and Member State consumer-protection authorities, including the Dutch Authority for Consumers & Markets guidelines on transparent and fair online-platform practices.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, the risk to the Right to Consumer Protection is assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of significant harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

	Lowest likelihood	Medium likelihood	Highest likelihood

[redacted]

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. The framework operates principally through the underlying Category 1 mitigation set (Sale of Prohibited Products in Section 4.1.3, Harmful False Information in Section 4.1.10, and Fraud and Spam in Section 4.1.11), the data-protection mitigation set (Section 4.2.4), and the advertising-side mitigations (Section 5.7). Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. The platform-architecture factors that constrain amplification operate to limit the structural opportunity for consumer-protection harm. Snap places specific emphasis on Article 25 DSA compliance in respect of online interface design, with internal review processes designed to prevent design patterns that could deceive or manipulate users or that materially distort users' free and informed decision-making. The default-settings dark-pattern review recorded in Section 5.1.3 provides additional review of design-pattern decisions affecting consumer-protection outcomes. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Harmful, False or Deceptive Practices and the Advertising Policies prohibit the categories of conduct most directly engaging consumer-protection rights. The Commercial Content Policy requires creators and partners to label paid or sponsored content clearly. Substantive enforcement detail is recorded in Section 4.1.10, Section 4.1.11 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. [redacted] The Article 16 DSA notice-and-action mechanism with the elevation of the illegal content reporting option (Section 5.4.6) supports the reactive-reporting dimension.

[redacted]

Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend deceptive content, scam content or fraudulent advertising; no corresponding interest categories exist. Steps are taken to limit surfacing of misleading or sensationalist headlines (Section 4.1.10). Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. The Advertising Policies, the pre-publication ad review pipeline [redacted] the prohibition on advertising-targeting based on minor users' activity data, the prominent ad markings supporting Article 26 DSA transparency, and the discontinuation of EU political advertising as of October 2025 together constitute the principal advertising-side mitigations. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Privacy-operations request volumes and advertising review, rejection, reporting and removal volumes are monitored on a continuing basis. Between 1 July 2025 and 30 June 2026, approximately [redacted] distinct advertisements completed final review and approximately [redacted] were rejected. Snap's H2 2025 global Transparency Report separately recorded 81,976 advertisements reported following publication, of which 47,266 were removed. Further detail is provided in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Limited application. Article 22 DSA trusted-flagger mechanisms in adjacent categories (Fraud and Spam, Section 4.1.11; Harmful False Information, Section 4.1.10) support the consumer-protection framework. Further detail in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Limited application. No specific code of conduct under Article 45 DSA is currently in operation in respect of consumer protection on Snapchat's in-scope services. Snap nevertheless cooperates with other providers and stakeholders through industry organisations and forums, including [redacted] on issues relevant to consumer protection, privacy and responsible platform governance. Snap also engages with EU and Member State authorities, including in connection with the [redacted] work on online financial

[redacted]

	misconduct recorded in Section 4.1.11. Further detail is provided in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Snap Privacy Policy, the Privacy Center, the Ads Privacy & Transparency Center, the Community Guidelines, the Harmful, False or Deceptive Information category explainer, the Safety Center, and Snap's Transparency Reports together provide consumer-relevant transparency. The Article 26 DSA advertisement-transparency requirements (clear identification of advertisements, the natural or legal person on whose behalf advertisements are presented, and the main targeting parameters) apply uniformly across the in-scope advertising surfaces. Further detail in Section 5.3 and Section 5.7.
Protection of Minors - Art. 35(1)(j)	Applies. Specific minor-protection measures bearing on consumer protection include: the prohibition on advertising-targeting based on minor users' activity data; [redacted]; the contact-restriction defaults; Family Center oversight including content controls; and minor-specific JIT notices warning of contact from accounts [redacted]. Further detail in Section 4.4.3 and Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. The Community Guidelines apply without regard to whether deceptive consumer-protection-relevant content is user-generated or AI-generated, including in respect of synthetic content depicting public figures or celebrities for fraudulent endorsement purposes. Generative-AI features are subject to the content-authenticity measures described in Section 5.9 (watermarking, contextual identification of generative-AI content). Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place in respect of the Right to Consumer Protection on Snapchat's in-scope services are reasonable, proportionate and effective.

The framework operates principally through the underlying Category 1 mitigation sets addressing Sale of Prohibited Products, Harmful False Information, and Fraud and Spam; the data-protection framework; and the advertising-side mitigations recorded in Section 5.7. These measures include

pre-publication review of all advertising, post-publication reporting and enforcement, Article 25 DSA interface-design controls, Article 26 DSA advertisement-transparency requirements, audience-floor requirements, special-targeting controls, restrictions protecting minor users, and the discontinuation of EU political advertising as of October 2025.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be monitored through the Section 6 ongoing risk-detection processes, with particular attention to evolving fraud and deceptive-advertising techniques, generative-AI-enabled misconduct, and the operational effectiveness of the advertising review and reporting systems.

4.2.8 Right to Property

This Section 4.2.8 sets out Snap's assessment of the actual or foreseeable risk of negative effects on the Right to Property arising from the operation of Snapchat's in-scope services. The Right to Property is recognised under Article 17 of the EU Charter (paragraph (1) addressing property generally, and paragraph (2) addressing intellectual property specifically) and under Article 1 of Protocol No. 1 to the European Convention on Human Rights.

Consistent with the 2025 Report, the substantive risk assessment, mitigation framework, and conclusion in respect of the Right to Property are recorded under the relevant Category 1 Sections rather than as a freestanding analysis in this Section. This Section accordingly serves as a structural cross-reference, ensuring comprehensive treatment of the Category 2 fundamental-rights risks while avoiding duplication of the substantive analysis.

Likelihood and severity

Two property dimensions are identified as bearing on Snapchat's in-scope services, each addressed in detail in the indicated Category 1 section.The principal property right potentially impacted by the operation of Snapchat's in-scope services is the right to intellectual property. The risk arises from the dissemination or use of content that infringes the intellectual property rights of a natural or legal person. The substantive risk assessment, mitigation framework, and conclusion in respect of this dimension are recorded in Section 4.1.5 (Dissemination of Content that Infringes on Intellectual Property Rights).

The other property right potentially impacted by the operation of Snapchat's in-scope services is the general right to property (physical property). The 2025 Report identified a residual potential risk that users may damage another person's physical property while creating content for viral or social distribution, with reference to reported instances on other social media platforms of property damage in connection with influencer activity and viral challenges. The substantive risk assessment, mitigation framework, and conclusion in respect of this dimension are recorded in Section 4.1.9 (Dissemination of Content Encouraging or Engaging in Violent or Dangerous Behaviour), where the relevant viral-challenge content sub-category is addressed.

[redacted]

Underlying Category 1 risk	2026 likelihood	2026 severity
Section 4.1.5 — Dissemination of Content that Infringes on Intellectual Property Rights	Lowest	Significant harm industry-wide
Section 4.1.9 — Dissemination of Content Encouraging or Engaging in Violent or Dangerous Behaviour	Lowest	Significant harm industry-wide

Likelihood and severity conclusion

On the basis of the substantive treatment recorded in Section 4.1.5 (Intellectual Property Infringement) and Section 4.1.9 (Violent or Dangerous Behaviour), the overall likelihood on Snapchat's in-scope services is assessed at the **lowest likelihood** category, and the severity as **Significant harm industry-wide**. Therefore the risk to the Right to Property is assessed at **Level 3** overall potential risk prioritisation in respect of both dimensions, consistent with the determinations recorded in those Sections.

The mitigations in place for the protection of the Right to Property are the same as those recorded in those Sections, mapped to Article 35(1)(a)–(k) DSA. There is no change in this overall determination from the 2025 Report.

[redacted]

Conclusion

Snap has concluded that the mitigations in place for the protection of the Right to Property on Snapchat's in-scope services are reasonable, proportionate and effective. The substantive basis for this conclusion is recorded in Section 4.1.5 (Intellectual Property Infringement) and Section 4.1.9 (Violent or Dangerous Behaviour). There is no change in this conclusion from the 2025

Report. Continued monitoring of both dimensions will operate through the Section 6 ongoing risk-detection processes as recorded in those underlying Sections.

4.3 Category 3 — Negative Effects on Democratic and Electoral Processes, Civic Discourse, and Public Security

This Section 4.3 sets out Snap's assessment of the actual or foreseeable negative effects of Snapchat's in-scope services on:

- democratic and electoral processes;
- civic discourse; and
- public security,

consistent with Article 34(1)(c) DSA and Recital 82.

In assessing each Category 3 risk, Snap considered the general risk factors set out in Article 34(2) DSA, in particular: (i) the extent to which the risk is influenced by intentional manipulation of the service, including by inauthentic use or automated exploitation, and by coordinated disinformation campaigns; (ii) the extent to which Snapchat's in-scope services allow for amplification and potentially rapid and wide dissemination of the relevant content; and (iii) specific regional or linguistic aspects, where applicable, including in connection with elections in individual Member States.

Summary of Category 3 determinations

The table below summarises Snap's assessment of likelihood, harm classification, overall potential risk prioritization, and the conclusion on mitigations under Article 35(1) DSA for each Category 3 risk. Consistent with Section 3.3, a risk is classified as severe harm industry-wide where it either: (i) presents a risk of significant damage to the physical or emotional well-being of Snapchatters or society at large; or (ii) presents an imminent, credible risk of severe harm, including threats to human life, safety or well-being. Overall potential risk prioritization is determined separately under the matrix and case-by-case framework recorded in Section 3.5.

Risk	Likelihood	Harm classification	Risk prioritization	Conclusion (Art. 35(1) DSA)
4.3.1 Negative Effects on Democratic and Electoral Processes	Lowest	Severe	**Level 3**	Reasonable, proportionate and effective
4.3.2 Negative Effects on Civic Discourse	Lowest	Severe	**Level 3**	Reasonable, proportionate and effective

[redacted]

4.3.3 Negative Effects on Public Security	Lowest	Serious	**Level 3**	Reasonable, proportionate and effective

Note: The determinations in the summary table reflect the outputs of the individual assessments in Sections 4.3.1 through 4.3.3 below. Where a determination has changed relative to the 2025 Report, that change is identified and explained in the corresponding section.

Each Category 3 risk is addressed in turn in the sub-sections that follow.

4.3.1 Negative Effects on Democratic and Electoral Processes

This Section 4.3.1 sets out Snap's assessment of the actual or foreseeable negative effects of Snapchat's in-scope services on democratic and electoral processes.

Harm Description

Negative effects on democratic and electoral processes may arise where content or activity on an online service misleads users about electoral procedures, discourages or obstructs lawful participation, promotes fraudulent or unlawful participation, or seeks to delegitimise an electoral or civic process. These effects may also arise through coordinated manipulation, deceptive synthetic media or the amplification of false or misleading claims concerning candidates, voting processes or election outcomes. This section assesses the extent to which those risks may arise on Snapchat's in-scope services and the measures in place to prevent, detect and respond to them.

Likelihood

Audience scale

Snapchat has a sizable European user base, with 96.5 million average monthly active recipients in the European Union during the 2026 assessment period. The 18–24 cohort comprised [redacted] recipients, or approximately [redacted] of the EU user population, while the 13–17 cohort comprised [redacted] recipients, or approximately [redacted] There is therefore, in principle, potential for public content on Snapchat to reach a material audience within voter-eligible age cohorts, noting that the voting age is 16 for certain elections in a number of Member States, including Austria, Belgium, Germany, Greece and Malta.

Platform architecture

Snapchat's in-scope services are configured in a way that, as a general matter, reduces the likelihood that political disinformation will achieve wide public amplification. Content is ephemeral by default; the service does not provide a broadcast feed for user-generated political content; political content is not eligible for algorithmic promotion in Spotlight; Discover political content is

limited to trusted news publishers and creators onboarded by Snap; and content achieving large-scale reach is subject to human moderation as a virality circuit-breaker.

Internal data

The assessment draws on the updated prevalence and enforcement data for Harmful False Information recorded in Section 4.1.10. Those data continue to indicate a low observed prevalence of violating content within that broader policy category. Because Harmful False Information includes content unrelated to democratic or electoral processes, those figures are treated as contextual evidence rather than as a direct measure of election-specific misinformation on Snapchat's in-scope services.

External evidence

External evidence specific to Snapchat's in-scope services is consistent with the low-likelihood determination:

- **Election Integrity Partnership (EIP)**. In connection with a major, high-profile election in 2022, Snap onboarded to the Election Integrity Partnership (EIP) — a coalition of research centers and civil society organisations that monitors online harms to democratic processes. As a participant in the EIP threat-escalation programme, Snap's teams received one incident report from the researchers monitoring risks on Snapchat during that election cycle.
- **European Commission stress test.** Ahead of the 2024 European Parliament elections, Snap participated in the European Commission's stress test and in multi-stakeholder roundtable dialogues, and was able to complete the stress-test exercises. As reported in Snap's 24 June 2024 election blog post, Snap observed a modest increase in reported activity but did not receive or observe material incidents or threats, and reported content was not verified as election misinformation on Snapchat.

Likelihood conclusion

Taking into account the platform architecture, external evidence and internal data above, the likelihood of a negative effect on democratic and electoral processes on Snapchat's in-scope services is assessed at the **lowest likelihood** category.

Severity

Negative effects on democratic and electoral processes are assessed as a risk of severe harm industry-wide. Disinformation targeting elections can be designed to undermine trust in electoral processes and to dissuade specific groups from exercising their voting rights. Industry-level examples, including the Cambridge Analytica matter, the 2022 Brazilian general election, and coverage of successive US presidential elections, illustrate the potential disruptive effects that digital platforms can have on democratic and electoral processes where safeguards are absent.

[redacted]

In light of this potential, the severity of harm is classified as **severe industry-wide**. This is unchanged from prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this harm on Snapchat's in-scope services. While the factors could, in principle, influence this systemic risk, their overall impact is limited in view of the mitigations described in Snap's Mitigations below and in Section 5.

Design of recommender systems and other algorithmic systems	Political content is not eligible for algorithmic distribution in Spotlight. On Discover, political and election-related content may be distributed from trusted news publishers and creators onboarded by Snap, subject to the applicable content and recommendation-eligibility policies. Spotlight and non-professional Discover content are subject to automated moderation and, [redacted] human review before wider recommendation. Further detail is provided in Section 5.6.
Content moderation systems	Proactive and reactive moderation procedures address content that undermines the integrity of elections or civic processes, [redacted] Moderators receive specific escalation guidance during elections. Policy coverage applies to all content formats, including AI-generated or modified media. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines prohibit content that undermines the integrity of elections and civic processes, including content under each of the four policy pillars identified above. Enforcement applies without regard to whether the content is user-generated or AI-generated. Further detail in Section 5.2.
Systems for selecting and presenting advertisements	As of October 2025, and consistent with the EU's Political Advertising Regulation, Snap has discontinued political advertising in the European Union. Advertising that remains in scope is subject to pre-publication review, automated and human moderation, and reactive reporting pathways. Snap's political-ad library continues to make available historical information on political advertising. Further detail in Section 5.7.
Data-related practices	Privacy-by-design and safety-by-design practices apply to product and feature changes that could bear on this risk, including changes to public-content distribution, recommendation systems and election-period safeguards. Relevant changes are reviewed

[redacted]

	through the Online Interface Design Process and, where applicable, the DDIA framework. Further detail is provided in Sections 5.1.5 and 6.3.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	In-scope services are configured to limit the rapid and wide amplification of content that may negatively affect democratic or electoral processes. Relevant design factors include short default retention for public stories (unless saved to Public Profile), the absence of broad reshare functionality of Stories, the non-eligibility of political content for algorithmic distribution in Spotlight, and pre-distribution moderation for public content recommended to wide audiences on Spotlight and Discover. On Discover, political and election-related content may be distributed from trusted news publishers and creators onboarded by Snap, subject to the applicable content and recommendation-eligibility policies.
Regional and linguistic aspects	The service is available in multiple EU languages, Terms are published in all EU languages, and moderators operate across multiple EU languages. Election-specific escalation guidance and partner engagement is calibrated to the Member State context in which an electoral event is taking place.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, negative effects on democratic and electoral processes are assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of severe harm industry-wide with lowest likelihood, without case-by-case deviation. The Level-3 classification reflects the low observed and foreseeable likelihood of material harm on Snapchat's in-scope services, rather than any diminution in the severity of the underlying risk. There is no change in this determination from the 2025 Report.

	Lowest likelihood	Medium likelihood	Highest likelihood

[redacted]

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function — Art. 35(1)(a)	Applies. Platform architecture limits the virality and public amplification of user-generated political content: content is ephemeral by default, the service does not provide a public broadcast feed for user-generated political content, and content reaching wide distribution on Spotlight is subject to moderation review. Further detail in Section 5.1.
Terms and Enforcement — Art. 35(1)(b)	Applies. The Community Guidelines prohibit content that undermines the integrity of elections and civic processes, organised around four policy pillars (procedural interference, participation interference, fraudulent or unlawful participation, and delegitimisation of civic processes). Influencer commercial-content policy prohibits election-related, advocacy and issue commercial content outside paid political advertising formats. Further detail in Section 5.2.
Moderation — Art. 35(1)(c)	Applies. [redacted] Further detail in Section 5.4.
Algorithmic Systems — Art. 35(1)(d)	Applies. [redacted] Further detail in Section 5.6.

[redacted]

Advertising Systems — Art. 35(1)(e)	Applies, with a material 2026 update: Snap discontinued political advertising in the European Union as of October 2025, consistent with the EU's Political Advertising Regulation. [redacted] Further detail in Section 5.7.
Ongoing Risk Detection and Management — Art. 35(1)(f)	Applies. Harmful False Information Prevalence Testing, relevant Transparency Report outputs, election-specific enforcement and escalation data, and external-partner reporting are used to monitor risks to democratic and electoral processes on a continuing basis. The PVP measure is treated as a broader contextual indicator rather than a direct election-specific prevalence measure. Further detail in Section 6.
Trusted Flaggers — Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA may concern content affecting democratic or electoral processes and are processed through the applicable priority-notice procedures. Election-specific concerns may also be escalated through the external-partner and crisis-response channels described in Sections 5.10 and 5.12. Further detail is provided in Section 5.10.
Codes and Crisis Protocols — Art. 35(1)(h)	Applies. Participation in industry coordination includes the Election Integrity Partnership (EIP); the AI Elections Accord; the Dutch Code of Conduct for political ads (2021); Logically Facts (fact-checking partnership ahead of the 2024 European Parliament elections); and the EU Code of Practice on Disinformation. Snap participated in the Commission's stress test and multi-stakeholder roundtable dialogues ahead of the European Parliament elections in 2024. Further detail in Section 5.12.
Transparency — Art. 35(1)(i)	Applies. Safety Center, blog posts and election-period communications provide contextual information about electoral integrity policies. The political-ad library remains available for historical disclosure. Where the service has partnered with electoral authorities for voter-information initiatives [redacted] those initiatives are published through in-app and blog channels. Further detail in Section 5.3.
Protection of Minors — Art. 35(1)(j)	Applies. Contact restrictions, Family Center, reporting flows and parent-facing guidance apply horizontally. For minors within the 16–17 voter cohort (in Member States where the voting age is 16), policy application is consistent with the general approach. Further detail in Section 5.8.

[redacted]

Content Authenticity — Art. 35(1)(k)	Partially applies. Community Guidelines apply without regard to whether content is user-generated or AI-generated, including deep fakes relating to elections. Contextual symbols identify Snap-generated AI content in certain surfaces. [redacted] Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for negative effects on democratic and electoral processes on Snapchat's in-scope services are reasonable, proportionate and effective.

The Commission's Guideline for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes (C/2024/3014) has been taken into account in the assessment. Responsive to that guidance, the mitigations identified include (i) the discontinuation of political advertising in the European Union, (ii) the non-eligibility of political content for algorithmic promotion in Spotlight, (iii) partnership with electoral authorities for voter-information initiatives, (iv) participation in the Election Integrity Partnership and the AI Elections Accord, and (v) the Community Guidelines coverage of content under the four policy pillars of electoral integrity.

There is no change in the overall conclusion from the 2025 Report.

4.3.2 Negative Effects on Civic Discourse

This Section 4.3.2 sets out Snap's assessment of the actual or foreseeable negative effects on civic discourse arising from the operation of Snapchat's in-scope services. Civic discourse is identified in Article 34(1)(c) DSA, alongside negative effects on democratic and electoral processes (Section 4.3.1) and public security (Section 4.3.3), as a systemic-risk category that VLOPs are required to assess and mitigate. This section should be read together with Section 4.3.1 (with which it shares substantial mitigation overlap) and Section 4.3.3.

Harm Description

Negative effects on civic discourse may arise where disinformation, misinformation, hateful or extreme content, or the amplification of such content degrades the quality of public debate, contributes to polarisation, or undermines the shared informational basis necessary for open democratic participation.

Likelihood

[redacted]

Underlying risk-pathway determinations

Negative effects on civic discourse are not captured through a standalone moderation or reporting category. The likelihood assessment therefore uses the determinations for Illegal Hate Speech, Harmful False Information, and Democratic and Electoral Processes as observable proxies. These risks capture the principal content and manipulation pathways through which Snapchat's in-scope services could negatively affect civic discourse. They are considered alongside any civic-discourse-specific reports and the relevant platform-design factors, rather than as a complete measure of the risk.

Underlying risk	2026 likelihood
Section 4.1.2 - Dissemination of Illegal Hate Speech	Lowest
Section 4.1.10 - Dissemination of Harmful False Information	Lowest
Section 4.3.1 - Negative Effects on Democratic and Electoral Processes	Lowest

The lowest-likelihood determination across all three underlying risks supports the determination that the volume of content presenting risks to civic discourse on Snapchat's in-scope services is low. The empirical observations recorded in those sections, in particular, [redacted] for Harmful False Information at the displayed precision, the very low VVR figures across surfaces and audiences for both Hate Speech and Harmful False Information, and the rapid median turnaround time for enforcement action against Harmful False Information (one minute), support this determination.

External evidence

External evidence relevant to civic-discourse risk on online platforms includes:

- **Cross-platform research on online-discourse dynamics.** A 2024 study analysing approximately 500 million comments across eight platforms over 34 years identified persistent patterns of toxic interaction across platforms, topics and time. Although toxic content was generally low in absolute terms, longer conversations were consistently associated with higher toxicity, and controversy was positively associated with toxicity in several datasets. The findings indicate that degradation of online discourse is a recurring, though not universal, risk in social-media environments. The study did not examine

Snapchat and is therefore used as general external evidence rather than evidence of incidence on Snapchat[77].

- **Academic and policy research on platform-level civic-discourse harms.** Research on echo chambers, algorithmic-amplification effects, and the role of platform architecture in shaping civic-discourse outcomes is well-developed across academic and policy literature. Research has identified features such as algorithmically curated feeds, frictionless resharing, ideologically segregated group structures, and engagement-based ranking as potential contributing factors. The limited applicability of these factors on Snapchat, including the absence of an unvetted political-content feed, broad reshare functionality of Stories and algorithmically promoted groups, operates as a structural mitigation in respect of civic-discourse risk.[78]

Likelihood conclusion

On the basis of the underlying risk-pathway determinations and proxy metrics for Hate Speech, Harmful False Information, and Democratic and Electoral Processes, together with the relevant platform-design factors, the likelihood of negative effects on civic discourse on Snapchat's in-scope services is assessed at the **lowest likelihood** category. There is no change in this determination from the 2025 Report.

Severity

Negative effects on civic discourse can have material consequences for democratic functioning, social cohesion, and the quality of public deliberation. Civic discourse is the connective tissue of democratic society, the public conversation through which citizens form opinions on shared concerns, exchange perspectives across differences, and hold institutions accountable. Where this discourse is degraded by polarisation, disinformation or extreme content amplification, the resulting harm can be both deep and durable, with consequences extending well beyond the immediate users affected.

Recognising the depth and durability of potential civic-discourse harm, the severity of negative effects on civic discourse is assessed as **severe harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

[77] Michele Avalle et al., "Persistent Interaction Patterns Across Social Media Platforms and Over Time", Nature, Vol. 628 (2024), pp. 582–589, doi:10.1038/s41586-024-07229-y.

[78] Philipp Lorenz-Spreen et al., "A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy", Nature Human Behaviour, Vol. 7 (2023), pp. 74–101, doi:10.1038/s41562-022-01460-1; S. Shyam Sundar et al., "Sharing Without Clicking on News in Social Media", Nature Human Behaviour, Vol. 9 (2025), pp. 156–168, doi:10.1038/s41562-024-02067-4; Jonas Stein, Marc Keuschnigg and Arnout van de Rijt, "Network Segregation and the Propagation of Misinformation", Scientific Reports, Vol. 13 (2023), Article 917, doi:10.1038/s41598-022-26913-5; William J. Brady et al., "Redesigning Algorithms to Intervene on Social Norm Misperceptions During a National Election", Nature, Vol. 655 (2026), pp. 942–956, doi:10.1038/s41586-026-10536-1.

[redacted]

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems do not, in the ordinary course, surface political or civic-discourse-engaging content as a category. Snapchat does not provide an unvetted feed of algorithmically curated political information. On Discover, political and civic content may be distributed from trusted news publishers and creators onboarded by Snap, subject to the applicable content and recommendation-eligibility policies. Political content is not eligible for algorithmic distribution in Spotlight, [redacted] The absence of an algorithmically-promoted-groups model is a structural feature, such groups have been identified in academic and policy research as contributing to echo chambers and as vectors for misinformation, with negative consequences for civic discourse. Further detail in Section 5.6.
Content moderation systems	Substantive moderation framework recorded in Section 4.1.2 (Hate Speech), Section 4.1.10 (Harmful False Information), and Section 5.4. [redacted] Snap engages with experts from across civil society and the research community who study information integrity and resilience to online harms.
Applicable terms and conditions and enforcement	The Community Guidelines on Hateful Content (Section 4.1.2) and the Harmful, False or Deceptive Information category (Section 4.1.10) prohibit content that engages civic discourse harms. The Community Guidelines apply uniformly across in-scope services and equally to all users. Substantive enforcement detail is recorded in Sections 4.1.2, 4.1.10 and 5.5.

[redacted]

Systems for selecting and presenting advertisements	Specific civic-discourse mitigations applicable to advertising are recorded in Section 4.3.1 and Section 5.7, and include: the discontinuation of EU political advertising as of October 2025 (consistent with Regulation (EU) 2024/900)[redacted]; the maintenance of the political advertising library for historical disclosure; the prohibition on advertising-targeting based on minor users' activity data; [redacted]
Data-related practices	PASS Review and DDIA framework (Section 4.2.4) apply to product and feature changes that may bear on civic discourse, including in respect of recommender system design and advertising targeting. Privacy by Design and the absence of profiling-based political ad targeting (consistent with Article 26(3) DSA prohibition on profiling using special-category personal data for advertising targeting) operate at the product-design level.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of civic-discourse-engaging content on in-scope services is structurally constrained by the platform architecture: recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences; the non-eligibility of political content for algorithmic distribution in Spotlight; and the technical barrier to Lens creation (which makes Lenses an unattractive vector for civic-discourse-related amplification compared to other distribution surfaces).

[redacted]

Regional and linguistic aspects	Member State election cycles, civic events and Member State-specific civic-discourse concerns are monitored individually by the Trust & Safety team, with the Content Crisis Response Protocol available for activation in elevated-risk situations. Engagement with Member State regulators [redacted]) is maintained on a continuing basis. ALD banned-word lists and misinformation classifiers operate across multiple EU languages.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, negative effects on civic discourse are assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of severe harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. The framework operates principally through the underlying Category 1 mitigation sets (Hate Speech in Section 4.1.2; Harmful False Information in Section 4.1.10) and the related Category 3 mitigations (Section 4.3.1, Democratic and Electoral Processes), with the platform-architecture factors (recommended content in public feed is moderated; short default retention for public stories (unless saved to Public Profile); no broad reshare functionality of Stories; no algorithmically-promoted groups; pre-distribution moderation) operating as structural mitigations. Descriptions are illustrative; detailed descriptions are provided in Section 5.

[redacted]

Snapchat Design and Function - Art. 35(1)(a)	Applies. The platform-architecture factors that mitigate civic-discourse harm include: the absence of an unvetted algorithmically-curated political content feed; the non-eligibility of political content for algorithmic distribution in Spotlight; the default retention ceiling for Public Stories and Snaps on Snap Map (unless saved to Public Profile), limiting the temporal arc of influence; the technical barrier to Lens creation that makes Lenses an unattractive vector for broad civic-discourse-related distribution; the absence of broad reshare functionality of Stories; and the absence of an algorithmically-promoted-groups model. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The substantive Community Guidelines prohibitions on Hateful Content (Section 4.1.2) and Harmful, False or Deceptive Information (Section 4.1.10) operate as the principal terms-and-enforcement framework engaging civic discourse harms. Snap takes steps to explain the policy approach to safeguarding civic-discourse information environments through the Community Guidelines and periodic blog posts. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. [redacted] Content approved for broader audiences must comply with the Community Guidelines and the Content Guidelines for Recommendation Eligibility. Reactive reporting is available through the in-app reporting flow with dedicated reporting reasons. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems mitigate risks to civic discourse, including through the absence of algorithmically-promoted groups (which have been shown to contribute to echo chambers and to be vectors for misinformation, with negative consequences for civic discourse). The Content Guidelines for Recommendation Eligibility set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution. Further detail in Section 5.6.

[redacted]

Advertising Systems - Art. 35(1)(e)	Applies. The discontinuation of EU political advertising as of October 2025 removes a documented risk pathway in respect of civic-discourse harm; the political advertising library continues to be made available for historical disclosure. [redacted]. The prohibition on advertising-targeting based on minor users' activity data, [redacted] and the prohibition on advertisements containing unsubstantiated or false claims operate alongside. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Hateful Content (Section 4.1.2) and for Harmful False Information (Section 4.1.10) provides the principal data inputs for monitoring this risk. Snap engages with experts from across civil society and the research community who study information integrity and resilience to online harms, including in connection with elevated-risk events (such as elections). The Content Crisis Response Protocol is available for activation in elevated-risk events affecting civic discourse. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Trusted-flagger relationships in adjacent categories (Hate Speech in Section 4.1.2; Harmful False Information in Section 4.1.10) support the civic-discourse framework.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap is a signatory of the EU Code of Conduct on Countering Illegal Hate Speech Online (revised 20 January 2025), which bears on the hate-speech dimension of civic-discourse harm. The DSA Code of Conduct on Disinformation became operative under the DSA framework on 1 July 2025. Snap did not subscribe to the Code during the assessment period, consistent with the low prevalence of harmful false information recorded on Snapchat. Snap participated in the European Commission's stress test and multi-stakeholder roundtable dialogues ahead of the 2024 European Parliament elections. Further detail is provided in Section 5.12.

[redacted]

Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the substantive harm-category explainers (Hateful Content, Terrorism and Violent Extremism; Harmful, False or Deceptive Information), the Safety Center, the Privacy Center, Snap's Transparency Reports, and periodic blog posts on the policy approach to safeguarding civic-discourse information environments together provide transparency on the framework. Article 26 DSA advertisement-transparency requirements apply uniformly across the in-scope advertising surfaces. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Specific minor-protection measures include: the limited exposure of minor users to political content (consistent with the non-eligibility of political content for algorithmic distribution in Spotlight and the absence of EU political advertising as of October 2025); the surfacing of trusted news sources for current events and civic education; Family Center oversight; and the resources for parents and carers. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. The risk that generative-AI tools could be used to produce harmful civic-discourse content (including deepfakes, manipulated audio/video, and synthetic content depicting public figures) is recognised. The content-authenticity measures recorded in Section 5.9 (watermarking, contextual identification of generative-AI content, and the substantive prohibitions across the Community Guidelines applying without regard to content origin) operate against this risk pathway. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for negative effects on civic discourse on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); the non-eligibility of political content for algorithmic distribution in Spotlight; the absence of algorithmically-promoted groups; the technical barrier to Lens creation), the substantive moderation framework for Hate Speech (Section 4.1.2) and Harmful False Information (Section 4.1.10), the discontinuation of EU political advertising in October 2025, the audience-floor requirement preventing micro-targeting, the engagement with civil society and academic experts on information integrity, the Member State-level engagement [redacted] the EU Code of Conduct on Countering Illegal Hate Speech Online signatory status, and the Content

Crisis Response Protocol for elevated-risk events together support the conclusion that the mitigation framework operates effectively against this risk.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to elevated-risk civic events and the continued evolution of generative-AI-related civic-discourse risk patterns.

4.3.3 Negative Effects on Public Security

This Section 4.3.3 sets out Snap's assessment of the actual or foreseeable negative effects on public security arising from the operation of Snapchat's in-scope services. Public security is identified in Article 34(1)(c) DSA, alongside negative effects on democratic and electoral processes (Section 4.3.1) and civic discourse (Section 4.3.2), as a systemic-risk category that VLOPs are required to assess and mitigate.

Harm Description

Negative effects on public security may arise where content incites, facilitates or glorifies terrorism, violence or other threats to life and safety, or where harmful false information contributes to public disorder or interferes with effective responses to elevated-risk events.

The substantive public-security mitigation framework, the Community Guidelines prohibitions across the four underlying Category 1 categories, the Proactive Safety Detection System for severe-harm content, the Article 9 and 10 DSA pathways for orders to act against illegal content and orders to provide information, the Law Enforcement Operations team, the EU Internet Forum participation, the EU Code of Conduct on Countering Illegal Hate Speech Online signatory status, and the Content Crisis Response Protocol, remain unchanged.

Likelihood

Underlying risk-pathway determinations

Negative effects on public security are not maintained as a standalone moderation or reporting category. The likelihood assessment therefore uses the determinations and metrics for Illegal Hate Speech, Terrorist Content, Violent or Dangerous Behaviour, and Harmful False Information as observable proxies. These categories capture the principal content pathways through which Snapchat's in-scope services could negatively affect public security, including incitement, terrorist or violent content, threats to life and safety, and false information capable of contributing to public disorder or interfering with responses to elevated-risk events. They are considered together with public-security-specific operational signals and the relevant platform-design factors, rather than as a complete standalone measure of the risk.

[redacted]

Underlying risk	2026 likelihood
Section 4.1.2 - Illegal Hate Speech	Lowest
Section 4.1.4 - Terrorist Content	Lowest
Section 4.1.9 - Violent or Dangerous Behaviour	Lowest
Section 4.1.10 - Harmful False Information	Lowest

All four underlying risks are assessed at the lowest likelihood. Available H1 2026 Prevalence Testing results recorded PVP of [redacted] for Hateful Content, [redacted] for Terrorism and Violent Extremism Content at the displayed precision, [redacted] for Violent or Disturbing Content, [redacted] for Dangerous Activities, and [redacted] for Harmful False Information at the displayed precision. Together with the metrics recorded in the underlying sections, these results support the determination that the volume of content presenting risks to public security on Snapchat's in-scope services is low.

External evidence

External evidence demonstrates that the identified public-security risk pathways have materialised on online platforms in practice, including through the amplification of hateful or false content and coordination around elevated-risk events:

- amplification of hate-content against Rohingya communities in Myanmar, contributing to mass-violence outcomes;[79]
- disorder following the Southport attack in July 2024, where false claims concerning the attacker spread rapidly online and social media operated as a significant organising and amplification channel;[80]
- the experience of school shootings, with documented severe negative effects on surviving youth, including long-lasting consequences for mental health, educational and economic trajectories;[81]

[79] Amnesty International, The Social Atrocity: Meta and the Right to Remedy for the Rohingya, Index No. ASA 16/5933/2022, 29 September 2022.
[80] House of Commons Home Affairs Committee, Police Response to the 2024 Summer Disorder, Second Report of Session 2024–25, HC 381, 14 April 2025, paras 7–13 and 26–28.
[81] Maya Rossin-Slater et al., "Local Exposure to School Shootings and Youth Antidepressant Use", Proceedings of the National Academy of Sciences, Vol. 117, No. 38 (2020), pp. 23484–23489, doi:10.1073/pnas.2000804117; Marika Cabral et al., "Trauma at School: The Impacts of Shootings on Students' Human Capital and Economic Outcomes", Review of Economic Studies, Vol. 93, No. 1 (2026), pp. 327–365, doi:10.1093/restud/rdaf027.

- Capitol-related events of 6 January 2021, with documented contributions from social-media platform amplification of false election-fraud claims.[82]

These examples are industry-wide and not Snapchat-specific. They demonstrate that the identified public-security risk pathways have materialised on online platforms in practice and illustrate the potentially serious consequences where effective mitigations are absent.

Likelihood conclusion

On the basis of the underlying risk-pathway determinations, the supporting empirical observations recorded in those sections, and the platform-design factors described below, including the absence of an open recommender feed and broad reshare functionality of Stories, short default retention for public stories (unless saved to Public Profile), and pre-distribution moderation for content reaching wide audiences, the likelihood of negative effects on public security on Snapchat's in-scope services is assessed at the **lowest likelihood** category.

Severity

The severity of harm associated with negative effects on public security can be substantial. The underlying severity classifications range across significant and serious:

- Hate Speech (Section 4.1.2): Significant harm industry-wide;
- Terrorist Content (Section 4.1.4): Serious harm industry-wide;
- Violent or Dangerous Behaviour (Section 4.1.9): Significant harm industry-wide; and
- Harmful False Information (Section 4.1.10): Significant harm industry-wide.

Taking the higher of these (serious harm industry-wide, from Terrorist Content) as the severity reference point for this risk, and recognising the broader public-security consequences that platform-mediated content can produce in elevated-risk events (terror attacks, mass-violence events, riots, election-related incidents), the severity of negative effects on public security is assessed as **serious harm industry-wide**, consistent with prior Reports.

As recorded in Section 3.5 and consistent with the Category 3 introduction, where any specific instance presents an imminent, credible risk of severe harm to human life, safety or well-being, or has a material impact on society at large, that specific instance is treated as a Level 1 overall risk and handled accordingly through the prioritised-response and law-enforcement-escalation pathways described in Section 5.5.

DSA Risk Factors

[82] Padinjaredath Suresh Vishnuprasad et al., "Tracking Fringe and Coordinated Activity on Twitter Leading Up to the US Capitol Attack", Proceedings of the International AAAI Conference on Web and Social Media, Vol. 18, No. 1 (2024), pp. 1557–1570, doi:10.1609/icwsm.v18i1.31409; U.S. House of Representatives, Select Committee to Investigate the January 6th Attack on the United States Capitol, Final Report, H. Rep. 117-663, 22 December 2022.

[redacted]

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services. The substantive treatment of the risk factors as they bear on the underlying at 1 Categories is recorded in Sections 4.1.2, 4.1.4, 4.1.9 and 4.1.10; the treatment below records the public-security-specific dimensions.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend content presenting risks to public security; no corresponding interest categories exist. The platform-architecture factors that constrain amplification (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences) operate to limit the structural opportunity for content presenting risks to public security to be amplified at scale or distributed at velocity. Substantive detail in Sections 4.1.2, 4.1.4, 4.1.9, 4.1.10 and Section 5.6.
Content moderation systems	The proactive and reactive moderation framework operating against the underlying Category 1 categories (Hate Speech in Section 4.1.2; Terrorist Content in Section 4.1.4; Violent or Dangerous Behaviour in Section 4.1.9; Harmful False Information in Section 4.1.10) is the principal operational pathway for content engaging public security risk. The Proactive Safety Detection System for severe-harm content [redacted] a specific operational feature. Substantive detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Hateful Content; Terrorism and Violent Extremism; Threats, Violence and Harm; and Harmful, False or Deceptive Information together constitute the substantive terms-and-enforcement framework for content engaging public security risk. Severe-harm sub-categories trigger swift enforcement consequences and, where appropriate, proactive escalation to law enforcement. Substantive detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising content that engages the substantive prohibitions in the underlying Category 1 risks (hate, terrorism, violence, false information). Pre-publication review applies to all advertising. The discontinuation of EU political advertising as of October 2025 reduces the advertising-side risk pathway for content engaging in public-security-related civic-discourse harm. Substantive detail in Section 5.7.

[redacted]

Data-related practices	PASS Review and DDIA framework apply to product and feature changes that may bear on this risk. Snap's information-security controls and incident-response framework address cybersecurity-related public-security risks where applicable to the in-scope services.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	The platform-architecture factors that constrain amplification (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences; the non-eligibility of political content for algorithmic distribution in Spotlight; the technical barrier to Lens creation; the absence of an algorithmically-promoted-groups model) operate to limit the structural opportunity for content engaging public security risk to spread at scale or velocity. Where elevated-risk events occur, the Content Crisis Response Protocol supports rapid operational response. The H2 2025 global Transparency Report recorded median periods of one minute from detection to final enforcement action for both Threats and Violence and Harmful False Information - both supportive of the rapid-response dimension of public-security risk management.

[redacted]

Regional and linguistic aspects	ALD banned-word lists for content engaging public security risk are maintained across multiple EU languages and calibrated to Member State legal contexts where required. The Article 9 / 10 DSA pathway for receipt and processing of orders to act against illegal content from EU and Member State authorities supports cooperation with Member State public-security authorities. The Law Enforcement Operations team operates as the principal operational point of contact for Member State law enforcement. Engagement with Member State internet referral units is maintained on a continuing basis. [redacted] Substantive detail in Sections 4.1.2, 4.1.4, and Section 5.5.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, negative effects on public security are assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of serious harm industry-wide with lowest likelihood, without case-by-case deviation. There is no change in this determination from the 2025 Report.

[redacted]

Snap's Mitigations

[redacted]

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. The framework operates principally through the underlying Category 1 mitigation sets (Sections 4.1.2, 4.1.4, 4.1.9 and 4.1.10), with the platform-architecture factors and the Content Crisis Response Protocol providing the structural and elevated-risk-event mitigations. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. Snapchat is not, in the ordinary course, an attractive platform for distributing content engaging public security risk: it is difficult to reach a broad audience without going through pre-distribution moderation; the absence of a broad reshare functionality of Stories and the short default retention for public stories (unless saved to Public Profile) limit the temporal arc of influence; Spotlight (the broadcast platform for user-generated content) is proactively and pre-moderated before content can reach a wide audience; Discover features content from approved media publishers and content creators. Substantive detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Hateful Content; Terrorism and Violent Extremism; Threats, Violence and Harm; and Harmful, False or Deceptive Information together constitute the substantive terms-and-enforcement framework. Severe-harm sub-categories (including credible imminent threats to human life or to severe physical safety) trigger swift enforcement consequences and proactive escalation to law enforcement where appropriate. Substantive detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Substantive moderation framework recorded in Section 4.1.2, Section 4.1.4, Section 4.1.9, Section 4.1.10 and Section 5.4. The Proactive Safety Detection System operates against severe-harm content [redacted] Article 9 / 10 DSA orders from EU and Member State authorities are processed through dedicated operational pathways.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend content engaging public security risk; no corresponding interest categories exist. The Content Guidelines for Recommendation Eligibility set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution. Substantive detail in Section 5.6.

[redacted]

Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit content engaging the substantive Category 1 prohibitions.[redacted] The discontinuation of EU political advertising as of October 2025 reduces the advertising-side pathway for public-security-related civic-discourse harm. Substantive detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing for Hate Speech (Section 4.1.2), Terrorist Content (Section 4.1.4), Violent or Dangerous Behaviour (Section 4.1.9) and Harmful False Information (Section 4.1.10) provides the principal data inputs for monitoring this risk. Snap engages with cross-industry partners and with Member State authorities on emerging public-security risk patterns. Substantive detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Article 22 DSA trusted-flagger relationships in adjacent categories [redacted] support the public-security framework.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap participates in the EU Internet Forum (EUIF), which addresses, among other matters, terrorist and violent extremist content with public-security implications. Snap is a signatory of the EU Code of Conduct on Countering Illegal Hate Speech Online (revised 20 January 2025). The Content Crisis Response Protocol operates as the principal internal mechanism for elevated-risk public-security events, including in connection with mass-casualty events, terrorist incidents, large-scale violent unrest, and election-related incidents. Cooperation with Europol's EU Crisis Protocol arrangements is maintained through cross-industry channels. Further detail is provided in Section 5.12.

Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the substantive harm-category explainers, the Severe Harms explainer, the Safety Center, Snap's Transparency Reports (including median-turnaround-time figures for the underlying Category 1 categories), and periodic blog posts on safety-related operational matters together provide transparency on the framework. Substantive detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Specific minor-protection measures bearing on public security include: the contact-restriction defaults reducing the contact pathway for content engaging public security risk directed at minors; Family Center oversight including content controls; the filtering of sensitive content from minor-user content feeds; the in-app educational resources of Here for You for users encountering distress related to violence, terrorism or other public-security-relevant events. Substantive detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. The risk that generative-AI tools could be used to produce synthetic content engaging public security risk (deepfake threat content, manipulated incident footage, AI-generated terror-related content) is recognised. The content-authenticity measures recorded in Section 5.9 (watermarking, contextual identification of generative-AI content, and the Community Guidelines applying without regard to content origin) operate against this risk pathway. Substantive detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for negative effects on public security on Snapchat's in-scope services are reasonable, proportionate and effective.

The combination of the platform architecture (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation; Spotlight pre-moderation; Discover content from approved publishers and content creators), the substantive moderation framework across the four underlying Category 1 categories, the rapid median turnaround times for enforcement (1 minute for Threats and Violence and 1 minute for Harmful False Information - the fastest across all Community Guidelines categories), the Proactive Safety Detection System for severe-harm content with law-enforcement escalation pathways, the Article 9 / 10 DSA pathway for receipt of orders from EU and Member State authorities, the EU Internet Forum participation, the EU Code

of Conduct on Countering Illegal Hate Speech Online signatory status, and the Content Crisis Response Protocol for elevated-risk events together support the conclusion that the mitigation framework operates effectively against this risk.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to: the operational outputs of any 2026 elevated-risk events activating the Content Crisis Response Protocol; the operational outputs of cross-industry engagement on emerging public-security risk patterns; and the continued evolution of generative-AI-related public-security risk patterns.

4.4 Category 4 - Negative Effects on Public Health

This Section 4.4 sets out Snap's assessment of the actual or foreseeable negative effects of Snapchat's in-scope services within the scope of Article 34(1)(d) DSA. The assessment covers:

- **public health** (Section 4.4.1);
- **gender-based violence** (Section 4.4.2);
- **minors** (Section 4.4.3); and
- **serious negative consequences for the physical and mental well-being of the person, including addictions** (Section 4.4.4).

Consistent with Recital 83, the Category 4 assessment considers risks relating to the design, functioning and use of in-scope services, including through manipulation, for example through coordinated disinformation campaigns relating to public health, or through online interface design that may stimulate behavioural addictions of recipients of the service.

Where these refinements bear on a specific Category 4 risk, they are reflected in the relevant sections 4.4.1 through 4.4.4. They do not, as a general matter, alter the Level-1 to Level-3 prioritization determinations reached in prior Reports, but they are intended to support a transparent Article 35(1) proportionality analysis.

Summary of Category 4 determinations

The table below summarises Snap's assessment of likelihood, harm classification, overall potential risk prioritization, and the conclusion on mitigations under Article 35(1) DSA for each Category 4 risk. The matrix in Section 3.5 is applied as a guide. Where the circumstances of a particular risk warrant a higher prioritization, including because of the vulnerability of the affected population or the potential for an imminent, credible risk of severe harm, Snap may apply a case-by-case upward adjustment from the matrix default. Any such adjustment is identified and explained in the relevant section.

Risk	Likelihood	Harm classification	Risk prioritization	Conclusion (Art. 35(1) DSA)
4.4.1 Negative Effects on Public Health	Lowest	Severe	**Level 3**	Reasonable, proportionate and effective
4.4.2 Negative Effects on Gender-Based Violence	Lowest	Serious	**Level 2***	Reasonable, proportionate and effective
4.4.3 Negative Effects on Minors	Lowest	Severe	**Level 1***	Reasonable, proportionate mitigation measures; EC Article 28 guidance taken into account
4.4.4 Serious Negative Consequences on Physical and Mental Well-Being	Lowest	Severe	**Level 1***	Reasonable, proportionate and effective

*Indicates a case-by-case upward adjustment from the default Section 3.5 matrix output. The basis for each adjustment is explained in the relevant section.

Note on the Minors assessment

Likelihood for Section 4.4.3 (Negative Effects on Minors) is recorded as "Lowest". Section 4.4.3 assesses the identified underlying risk pathways separately and concludes that each falls within the lowest likelihood category. The overall Level 1 prioritisation reflects a case-by-case upward adjustment from the default matrix output, as explained in that section.

The conclusion for the Minors assessment records that reasonable, proportionate mitigation measures have been implemented, and that the European Commission's guidance on Article 28 has been taken into account in the assessment and continues to inform the approach. That framing is preserved in this Report and is further developed in Section 4.4.3.

Each Category 4 risk is addressed in turn in the sub-sections that follow.

4.4.1 Negative Effects on Public Health

This Section 4.4.1 sets out Snap's assessment of the actual or foreseeable negative effects of Snapchat's in-scope services on public health.

Without mitigation, digital content platforms can contribute to negative effects on public health. In respect of Snapchat's in-scope services, the risks considered in this section include, without limitation:

[redacted]

- glorification of self-harm, including the promotion of self-injury, suicide or eating disorders;
- health misinformation, including content relating to unproven treatments, supplements or other interventions not recognised or approved by the medical community;
- impacts on mental health, including disordered eating, anxiety, depression, and patterns associated with addictive engagement, in particular in respect of minors; and
- promotion of products that can cause health-related harm, including drugs and regulated products such as e-cigarettes (vapes).

Likelihood

Underlying risk-pathway determinations

The likelihood analysis for negative effects on public health is informed by the likelihood determinations for the underlying Category 1 risks that bear on this public-health risk, drawn from Section 4.1:

Underlying Category 1 risk	Likelihood
4.1.3 Sale of Prohibited Products (drugs, regulated products)	**Lowest**
4.1.8 Self-Harm, Suicide and Eating Disorders	**Lowest**
4.1.9 Violent or Dangerous Behaviour	**Lowest**
4.1.10 Harmful False Information (including health misinformation)	**Lowest**

Internal data

Prevalence Testing for the underlying content categories most closely aligned with public-health harms recorded the following results for the 90-day period ending 30 June 2026:

- Harmful False Information, including health misinformation; [redacted] at the displayed precision;
- Self-Harm and Suicide content: [redacted];
- Drugs: [redacted];
- Regulated Goods & Activities: [redacted]; and
- Illegal Goods & Activities; [redacted]

Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP figures were [redacted] for Harmful False Information at the displayed precision, [redacted] for Self-Harm and Suicide, [redacted] for Drugs, [redacted] for Regulated Goods & Activities, and [redacted] for Illegal Goods & Activities at the displayed precision. These refined

figures are presented as supplementary 2026 context and are not used to determine the year-on-year direction.

These categories are separately classified within the PVP taxonomy and should not be aggregated.

External Evidence

The external evidence described in the severity analysis below also supports the foreseeability of the public-health pathways assessed in this section, including health misinformation, eating-disorder and body-image harms, and youth exposure to regulated products such as e-cigarettes. That evidence is industry-wide and does not provide a Snapchat-specific prevalence estimate. The relative-likelihood determination therefore remains based principally on the underlying Category 1 determinations and the internal PVP evidence set out above.

Likelihood conclusion

On the basis of the underlying Category 1 determinations and the internal data above, the likelihood of a negative effect on public health on Snapchat's in-scope services is assessed at the **lowest likelihood** category.

Severity

Negative effects on public health are assessed as a risk of severe harm industry-wide. Industry-level evidence bearing on severity includes:

- During the COVID-19 pandemic, the World Health Organization identified the large-scale circulation of misinformation through digital and social-media channels as an "infodemic" that undermined the public-health response and jeopardised measures intended to control the pandemic. WHO separately documented false claims concerning preventative measures, treatments and purported cures[83].
- External evidence indicates that social-media content may contribute to body dissatisfaction and disordered eating. The National Eating Disorders Association states that eating disorders do not have a single cause, while media exposure can contribute to body dissatisfaction and disordered eating. NEDA's 2026 community survey also reported widespread triggering or escalation of eating-disorder thoughts and behaviours following exposure to body-focused content.[84]
- A study requested by the European Parliament's Committee on Culture and Education identified children's exposure to content promoting eating disorders as an online risk and

[83] World Health Organization, "Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation," 23 September 2020; World Health Organization, "Fighting misinformation in the time of COVID-19, one click at a time," 27 April 2021.
[84] National Eating Disorders Association, "Media and Eating Disorders"; National Eating Disorders Association, Social Media, Eating Disorders, and Self-Harm Risk: What Our Community Told Us (2026).

associated such exposure with unhealthy eating habits, body-image dissatisfaction, and distorted values and attitudes.[85]

- Research by the Dutch National Institute for Public Health and the Environment identifies vaping as harmful to health and flavoured e-cigarettes as particularly appealing to young people. The Netherlands introduced restrictions on e-cigarette flavours to reduce their attractiveness to young people and non-smokers.[86]

In light of this evidence, the severity of negative effects on public health is classified as **severe harm industry-wide**. This is unchanged from prior Reports.

DSA Risk Factors

The service and general risk factors set out in Article 34(2) DSA have been considered in the context of negative effects on public health. Because these factors, as a general matter, operate similarly across the underlying Category 1 harms and are addressed through horizontal mitigation measures that cut across multiple systemic risks, the detailed factor-by-factor analysis is not repeated in this section. The reader is directed to the Article 34(2) analyses in the corresponding Category 1 Sections: Section 4.1.3 (Sale of Prohibited Products or Services), Section 4.1.8 (Self-Harm, Suicide and Eating Disorders), Section 4.1.9 (Violent or Dangerous Behaviour), and Section 4.1.10 (Harmful False Information).

Section 4.1.3, including the feature-level assessment of Search and Find Friends as pathways for information relating to the sale of prohibited products, and updates to relevant moderation and interface measures, including Heads Up, are reflected in this public-health analysis. They do not alter the overall Article 34(2) factor analysis for Public Health.

Overall Potential Risk Prioritization

Applying the matrix in Section 3.5, and consistent with prior Reports, negative effects on public health are assessed at **Level 3** overall potential risk prioritization. This output corresponds to the matrix intersection of severe harm industry-wide with lowest likelihood, without case-by-case deviation. The Level-3 classification reflects the low observed and foreseeable likelihood of material harm on Snapchat's in-scope services, rather than any diminution in the severity of the underlying risk. As a general matter, and consistent with the approach recorded in Section 3.5, any specific issue that arises within this category and presents an imminent, credible risk of severe harm - including threats to human life, safety and well-being - is treated as a Level 1 overall risk and handled accordingly. There is no change in the Level-3 classification from the 2025 Report.

[85] European Parliament, Directorate-General for Internal Policies, The Influence of Social Media on the Development of Children and Young People, study requested by the Committee on Culture and Education (2023).
[86] Dutch National Institute for Public Health and the Environment (RIVM), Reducing the Attractiveness of E-Liquids to Youth: A Proposal for a Restrictive List of Tobacco Flavourings; RIVM, "Two in five reduced the use of e-cigarettes after the introduction of the flavour ban" (2025).

[redacted]

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA in the table below. Because the mitigation set for negative effects on public health is substantially co-extensive with the mitigations for the underlying Category 1 risks, the descriptions are illustrative; detailed descriptions are provided in Section 5, cross-referenced by sub-section.

Snapchat Design and Function - Art. 35(1)(a)	Applies. In-scope services are configured to include proactive moderation for content that can produce negative effects on public health, including content relating to prohibited or regulated products, self-harm, suicide, eating disorders, violent or dangerous behaviour, and harmful false information. In-app search for sensitive terms surfaces safety resources, including Here For You for suicide-related search terms. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines prohibit the glorification of self-harm, including the promotion of self-injury, suicide or eating disorders; harmful false or deceptive information; and the sale or promotion of prohibited or regulated products. Further detail in Section 5.2.
Moderation - Art. 35(1)(c)	Applies. Proactive and reactive moderation operates across in-scope public surfaces for each of the underlying content types identified above. Further detail in Section 5.4.
Algorithmic Systems - Art. 35(1)(d)	Applies. Recommender systems are configured not to surface content in the categories identified above; no corresponding interest categories exist for such content. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising policies prohibit content that would cause negative effects on public health, including promotion of drugs, weapons and certain regulated products. Advertising relating to

[redacted]

	health-related products and services is subject to additional review requirements. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing, Transparency Report outputs and enforcement data are used to monitor risks to public health on a continuous basis. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Limited application. Trusted-flagger relationships relevant to the underlying public-health risks are addressed in the corresponding Category 1 risks. Notices from competent authorities and engagement with specialist organisations concerning illegal drugs and regulated products operate alongside those relationships. Further detail in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. [redacted] Partnership with NGOs, third-party providers and law-enforcement authorities continues in relation to the prevention of illegal sale of vape-related products. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. Guidance on terms, harms, moderation and enforcement is published in the Safety Center. In-app safety resources include Here For You for suicide-related search terms and, where applicable, interstitials on sensitive topics. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Measures for minors include contact restrictions, Family Center, reporting flows, and guidance for parents and carers via the Snapchat Family Safety Hub and the parents site. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Partially applies. Measures are in place to mitigate (i) the use of generative-AI tools within the service to create illegal or otherwise violating content, including health misinformation and deepfakes, and (ii) the dissemination on in-scope services of such content created using generative-AI tools on other platforms. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for negative effects on public

health on Snapchat's in-scope services are reasonable, proportionate and effective. There is no change in the overall conclusion from the 2025 Report.

4.4.2 Negative Effects on Gender-Based Violence

This Section 4.4.2 sets out Snap's assessment of the actual or foreseeable negative effects on gender-based violence (GBV) arising from the operation of Snapchat's in-scope services. Article 34(1)(d) DSA identifies negative effects on the protection of victims of gender-based violence, alongside negative effects on public health and protection of minors, as a Category 4 systemic-risk category. Snap's mitigations in this category address conduct that promotes, facilitates, or constitutes gender-based violence, with the substantive operational framework operating principally through the underlying Category 1 categories engaging GBV-relevant conduct.

Applying the matrix in Section 3.5, the combination of serious harm industry-wide and lowest likelihood would produce a Level 3 prioritization by default. This section considers whether a case-by-case upward adjustment is warranted in light of the nature, scale and durability of gender-based violence harms. The basis for the resulting determination is set out in the Overall Potential Risk Prioritization discussion below.

Likelihood

Underlying risk-pathway determinations

Snap does not track Gender-Based Violence as a separate category in its Transparency Reports or in Prevalence Testing. GBV-relevant conduct is captured within the scope of the following broader categories tracked by Snap, each of which is the subject of a dedicated Category 1 risk and the substantive moderation framework recorded in Section 5.4:

Underlying risk	2026 likelihood
Section 4.1.2 – Illegal Hate Speech	Lowest
Section 4.1.6 – Adult Sexual Content (including Sexual Crimes, NCII and sextortion)	Lowest
Section 4.1.7 – Harassment and Bullying	Lowest
Section 4.1.9 – Violent or Dangerous Behaviour	Lowest

[redacted]

Section 4.1.10 – Harmful False Information	Lowest

Internal data - underlying category PVP

Across the assessment cycles recorded in this and prior Reports, prevalence within the underlying categories has remained low, with an overall downward trajectory notwithstanding limited variation in individual categories.

[redacted]

Using the methodology retained for year-on-year comparability in Section 6.4, the results for the 90-day period ending 30 June 2026 show a further decrease in Hateful Content, Harassment, Dangerous Activities and Harmful False Information. Adult Sexual Content remained broadly stable at a low absolute level, decreasing slightly from [redacted] while Violent and Disturbing Content increased slightly from [redacted] Taken together, the underlying categories continue to record low prevalence, with an overall downward trajectory across the assessment cycles.

Applying the refined double-review methodology described in Section 6.4, the corresponding confirmed-violation PVP figures for the 90-day period ending 30 June 2026 were [redacted] for Hateful Content, [redacted] for Harassment, [redacted] for Adult Sexual Content, [redacted] for Violent and Disturbing Content, [redacted] for Dangerous Activities, and [redacted] for Harmful False Information at the displayed level of precision. These refined figures are presented as supplementary 2026 context and are not used to determine the year-on-year direction.

The H2 2025 Global Transparency Report recorded 907,230 reports of Hate Speech content, resulting in 335,547 enforcements and enforcement against 276,999 unique accounts. These figures cover the whole of Snapchat, including private surfaces outside the scope of this Report, and are therefore used only as contextual evidence that the relevant reporting and enforcement pathways are actively used, rather than as a direct measure of GBV prevalence on the in-scope services. Compared with H1 2025, report volume increased modestly, while enforcement volumes and unique accounts enforced remained broadly stable.

External evidence

External evidence on platform-level GBV risk is well-developed. Selected sources include:

- **EU-wide evidence on gender-based and cyber violence.** The 2026 EU Gender-Based Violence Survey published by FRA and EIGE found that approximately one in three women in the EU had experienced physical and/or sexual violence. The accompanying analysis identifies cyber violence, psychological violence and economic violence as widespread but frequently under-recognised forms of abuse, and records substantial under-reporting to law-enforcement authorities.[87]
- **Online violence, misogyny and political participation.** Research by the National Democratic Institute identifies online threats, harassment, hate speech and gendered disinformation as tactics that may silence politically active women, cause withdrawal from political and civic participation, and undermine freedom of expression, information integrity and democratic participation.[88]
- **EU legal recognition of cyber violence.** Directive (EU) 2024/1385 establishes minimum EU rules addressing violence against women and domestic violence, including the non-consensual sharing of intimate or manipulated material, cyberstalking, cyber-harassment and cyber-incitement to violence or hatred. The Directive recognises that such conduct may cause serious psychological or physical harm and requires Member States to transpose its provisions by 14 June 2027.[89]
- **Technology-facilitated GBV and generative AI.** UN Women has identified technology-facilitated gender-based violence as encompassing conduct including online harassment, impersonation, intimate-image abuse, sextortion and gendered disinformation. Its more recent work identifies artificial intelligence as increasing the scale

[87] European Union Agency for Fundamental Rights and European Institute for Gender Equality, EU Gender-Based Violence Survey - Evidence for Policy and Practice (2026); European Union Agency for Fundamental Rights, “EU-wide survey highlights scale of psychological, economic, and cyber violence against women,” 3 March 2026.
[88] National Democratic Institute, Interventions to End Online Violence Against Women in Politics (2022); National Democratic Institute, Submission to the United Nations Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression regarding Gendered Disinformation (7 July 2023); National Democratic Institute, Tweets That Chill: Analyzing Online Violence Against Women in Politics (2019).
[89] Directive (EU) 2024/1385 of the European Parliament and of the Council of 14 May 2024 on combating violence against women and domestic violence, OJ L 2024/1385, 24 May 2024, in particular Articles 5–8 and 49.

and accessibility of harms including sexual deepfakes, automated hate, impersonation and AI-assisted sextortion.[90]

- **Financial sextortion and organised offending.** Thorn's research into NCMEC CyberTipline data found that financial-sextortion reports predominantly concern boys and young men and reflect the emergence of organised offenders using technology to target and extort minors at scale. The research also identifies movement from initial-contact platforms to secondary messaging services, including encrypted services, as a recurring operational pattern.[91]
- **Gendered disinformation as a democratic and security risk.** A 2026 European Parliamentary Research Service briefing identifies AI-enabled gendered disinformation as contributing to a more hostile information environment for women and affecting their democratic participation online and offline. It also records the use of gendered disinformation by authoritarian and hostile actors to undermine women leaders, polarise societies and weaken confidence in democratic rights and institutions.[92]

Likelihood conclusion

On the basis of the likelihood determinations for the underlying Category 1 risks, the internal prevalence data across the relevant underlying categories, and the platform-architecture factors described below, the likelihood of negative effects on gender-based violence arising from Snapchat's in-scope services is assessed at the **lowest likelihood category.**

Severity

The risks of negative effects on gender-based violence stemming from online platforms are well-documented and are recognised across EU and Member State legal frameworks, civil-society research, and academic literature. The harms stemming from online gender-based violence are not limited to the digital realm: GBV-relevant content can produce direct psychological, social, professional and physical-safety harm to those affected, and can chill the exercise of political and civic rights with disproportionate impact on women and girls. Specific examples preserved from the 2025 Report include:

- the chilling effect of online misogyny on women's participation in political and civic life;
- the role of NCII and sextortion in producing severe psychological harm and (in some cases) self-harm to victims;
- the role of gender-based disinformation in undermining women's professional and political standing;

[90] UN Women, Technology-Facilitated Gender-Based Violence: Developing a Shared Research Agenda (September 2024); UN Women, How AI Is Exacerbating Technology-Facilitated Violence Against Women and Girls (November 2025).
[91] Thorn, Trends in Financial Sextortion: An Investigation of Sextortion Reports in NCMEC CyberTipline Data (June 2024).
[92] European Parliamentary Research Service, Women in the Age of AI-Enabled Disinformation (4 March 2026).

[redacted]

- the role of organised-crime networks targeting teenage boys with financial sextortion using real or AI-generated material;
- the increasing role of generative-AI technologies in producing technology-facilitated GBV (including deepfake NCII).

On these analytical bases, and consistent with the 2025 Report and with Directive 2024/1385/EU's recognition of cyber-violence as a serious form of violence against women, the severity of negative effects on gender-based violence is assessed as **serious harm industry-wide**. As a general matter, and consistent with Section 3.5, any specific instance presenting an imminent, credible threat to physical safety or to severe psychological well-being is treated as a Level 1 overall risk and handled accordingly through the prioritised-response and law-enforcement-escalation pathways.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems are configured not to recommend content engaging GBV harm; no corresponding interest categories exist. The platform-architecture factors that constrain amplification (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation) operate to limit the structural opportunity for GBV-relevant content distribution. Limited remix functionality (restricted to specific content types and applying short retention) further constrains the amplification pathway. Further detail in Section 5.6.
Content moderation systems	The proactive and reactive moderation framework operating against the underlying categories (Section 4.1.2 Hate Speech; Section 4.1.6 Adult Sexual Content / sextortion; Section 4.1.7 Harassment and Bullying; Section 4.1.9 Violent or Dangerous Behaviour; Section 4.1.10 Harmful False Information) is the principal operational pathway for GBV-relevant content. [redacted] Further detail in Section 5.4.

[redacted]

Applicable terms and conditions and enforcement	The Community Guidelines on Hateful Content, on Harassment and Bullying, on Sexual Content (including the prohibition on NCII), on Threats, Violence and Harm, and on Harmful, False or Deceptive Information together prohibit conduct engaging GBV harm. The Community Guidelines apply uniformly across in-scope services and equally to all users. Substantive enforcement detail is recorded in the relevant underlying Sections.
Systems for selecting and presenting advertisements	Advertising Policies prohibit advertising containing hate-related, sexual, threatening or otherwise GBV-relevant content. The discontinuation of EU political advertising as of October 2025 removes paid political advertising as a potential distribution pathway for gender-based political disinformation in the EU. The prohibition on advertising-targeting based on minor users' activity data [redacted]. Further detail in Section 5.7.
Data-related practices	PASS Review and DDIA framework apply to product and feature changes that may bear on GBV risk, including specific consideration of GBV in Lens design review (recorded in Section 5.1.5). Privacy-by-design practices, including data protection by design and by default under Article 25 GDPR, operate at the product-development level to limit the data-exposure pathways for GBV-relevant misuse.
Intentional manipulation (Article 34(2))	[redacted]

[redacted]

Amplification and rapid, wide dissemination	Amplification of GBV-relevant content is structurally constrained by the platform architecture as recorded above. The default retention ceiling for Public Stories and Snaps on Snap Map (unless saved to Public Profile) further limits the temporal arc of influence of content engaging this category. Where GBV-relevant content is identified (proactively or reactively), the median turnaround times for enforcement action across the underlying categories are rapid (recorded in detail in the relevant underlying Sections).
Regional and linguistic aspects	ALD banned-word lists for hate-related and harassment-related content are maintained across multiple EU languages and calibrated to regional context where required. Member State frameworks on GBV-relevant criminal conduct vary materially across the EU; engagement with Member State child-protection authorities, women's-rights organisations, and law-enforcement bodies is maintained on a continuing basis. [redacted] Notices submitted by designated trusted flaggers under Article 22 DSA are separately subject to the priority-treatment framework described in Section 5.10.

Overall Potential Risk Prioritization

Given the lowest likelihood category and given the serious harm classification, the application of the default matrix described in Section 3.5 would lead to a Level 3 risk prioritization. However, applying the case-by-case deviation framework recorded in Section 3.5 of the Methodology, this risk is assessed at **Level 2** overall potential risk prioritization. The deviation up from Level 3 to Level 2 is supported by the analytical bases recorded in the Severity discussion above and by the following considerations:

- the recognised global scale and durability of GBV harm, and the elevated seriousness with which the EU has approached the harm through Directive 2024/1385/EU and the Istanbul Convention;
- the disproportionate impact of GBV harm on women and girls, and the chilling effect on the exercise of political and civic rights documented across multiple research bases;
- the recognition of technology-facilitated GBV as an emerging risk dimension warranting elevated attention as generative-AI technologies develop;
- the operational interconnection of GBV-relevant conduct across multiple Category 1 underlying categories, supporting an integrated mitigation framework rather than a single-category response.

[redacted]

There is no change in this determination from the 2025 Report. On these bases, Snap applies a case-by-case upward adjustment from the Level 3 matrix default and assesses this risk at Level 2 overall potential risk prioritization.



Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)–(k) DSA below. The framework operates principally through the underlying Category 1 mitigation sets (Sections 4.1.2, 4.1.6, 4.1.7, 4.1.9, 4.1.10), supplemented by GBV-specific design considerations (Lens design review) and the cross-platform sextortion measures recorded in Section 4.1.6. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. Several aspects of Snapchat's design and function reduce the risk of GBV-relevant content being shared on the platform: pre-distribution moderation for content [redacted]; the absence of an open recommender feed for unvetted user-generated content; the limited prevalence design for sexually-suggestive content; short default retention for public stories (unless saved to Public Profile); the limited remix functionality restricted to specific content types with short retention. The Lens design framework includes specific GBV-related review (recorded in Section 5.1.5). Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Community Guidelines on Hateful Content, on Harassment and Bullying, on Sexual Content (including the explicit prohibition on NCII), on Threats, Violence and Harm, and on Harmful, False or Deceptive Information together prohibit conduct engaging GBV harm. The Sexual Crimes sub-section of the Adult Sexual Content category (Section 4.1.6) specifically addresses sextortion and NCII. Further detail in Section 5.2 and Section 5.5.

[redacted]

Moderation - Art. 35(1)(c)	Applies. Substantive moderation framework recorded in Sections 4.1.2, 4.1.6, 4.1.7, 4.1.9, 4.1.10 and Section 5.4. [redacted] Article 22 DSA trusted-flagger notices are separately handled under the priority-treatment framework described in Section 5.10.
Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend GBV-relevant content; no corresponding interest categories exist. Steps are taken to limit surfacing of misogynistic, harassing or sexually-violative content. The Content Guidelines for Recommendation Eligibility set additional standards for content eligible for algorithmic distribution. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Advertising Policies prohibit GBV-relevant content. Pre-publication review applies to all advertising. The discontinuation of EU political advertising as of October 2025 removes paid political advertising as a potential distribution pathway for gender-based political disinformation in the EU. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. Prevalence Testing across the underlying categories provides the principal data inputs for monitoring this risk. The results for the 90-day period ending 30 June 2026 continue to show low prevalence across the underlying categories, with further decreases in Hateful Content, Bullying and Harassment, Dangerous Activities and Harmful False Information, broadly stable Adult Sexual Content prevalence, and a slight increase in Violent and Disturbing Content. Snap engages with civil-society organisations focused on women's rights, with Member State child-protection authorities, and with law-enforcement bodies on a continuing basis.
Trusted Flaggers - Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA may concern GBV-relevant content and are processed under the priority-treatment framework described in Section 5.10. Separately, Snap's cooperation with Stop Fisha (France) supports the NCII-related dimension of GBV.

[redacted]

Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap is a signatory of the EU Code of Conduct on Countering Illegal Hate Speech Online (revised 20 January 2025), which bears on the hate-speech dimension of GBV. Snap engages with civil-society organisations focused on online violence against women and women's rights, with Member State authorities, and with cross-industry partnerships including (where relevant) the cross-platform sextortion measures recorded in Section 4.1.6. Snap engages with the Commission and Member States on the implementation of Directive 2024/1385/EU on combating violence against women. Further detail is provided in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the underlying-category explainers, the Safety Center, and Snap's Transparency Reports together provide transparency on the framework. Snap publishes resources on healthy relationships, consent, and digital wellness for teen users. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. Mitigations specific to minor users that bear on GBV include: (i) the contact-restriction defaults (Section 5.1.3) that limit unsolicited contact pathways used in [redacted] (ii) Family Center oversight; (iii) the dedicated "It involves a child" reporting reason; (iv) the prohibition on advertising-targeting based on minor users' activity data; and (v) the integration of the AADC Risk Assessment within the DDIA framework. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. The Community Guidelines apply without regard to whether GBV-relevant content is user-generated or AI-generated, including in respect of synthetic NCII (deepfake-driven GBV imagery). Generative-AI features are subject to the content-authenticity measures recorded in Section 5.9. The recognition of technology-facilitated GBV as an emerging risk dimension (recorded in the External evidence sub-section above) is acknowledged in the framework. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for negative effects on

[redacted]

gender-based violence on Snapchat's in-scope services are reasonable, proportionate and effective.

The framework operates through the integrated Category 1 mitigation sets covering the principal underlying categories (Hate Speech, Adult Sexual Content including sextortion and NCII, Harassment and Bullying, Violent or Dangerous Behaviour, Harmful False Information), supplemented by the platform-architecture factors that limit GBV-relevant content amplification (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation; limited remix functionality with short retention), the GBV-specific design considerations (Lens design review), the Article 22 DSA trusted-flagger relationship with [redacted] and the engagement with civil-society organisations, Member State authorities, and law-enforcement bodies on a continuing basis. The friending-protections, profile-level moderation expansion, and the discontinuation of EU political advertising further strengthen the framework on operational dimensions identified through Snap's continuing systemic-risk-assessment and mitigation processes.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to: the operational outputs of friending-protections; the operational outputs of the cross-platform sextortion measures recorded in Section 4.1.6; the implementation framework of Directive 2024/1385/EU as Member States transpose by the 14 June 2027 deadline; and the continued evolution of technology-facilitated GBV risk patterns as generative-AI capabilities develop.

4.4.3 Negative Effects on Minors

This Section 4.4.3 sets out Snap's assessment of the actual or foreseeable negative effects on minors arising from the operation of Snapchat's in-scope services. Article 34(1)(d) DSA identifies negative effects on the protection of minors, alongside negative effects on public health and gender-based violence, as a Category 4 systemic-risk category. This section, together with Section 4.2.6 (Children's Rights) and Section 4.4.4 (Physical and Mental Well-Being), constitutes the principal Report-wide treatment of risks engaging minor users.

Although the relative likelihood for the Negative Effects on Minors varies, Snapchat considers the risk of harm to fall within the severest category. Consequently, Snap considers this to be a **Level 1 overall potential risk**. There is no change in this assessment from our 2025 Report.

Scope of Section 4.4.3

[redacted]

Section 4.4.3 assesses the integrated risk to minor users arising from the operation of the in-scope services across all underlying substantive harm categories. The complementary sections are:

- **Section 4.2.6 (Children's Rights).** The fundamental-rights view of children's expression and digital participation, focused on the participation-and-expression element of Article 24 EU Charter and the corresponding UNCRC provisions.
- **Section 4.4.4 (Physical and Mental Well-Being).** The well-being-specific view of risks engaging minor populations, with concentration on the screen-time, online-pressure, and mental-health dimensions.

Read together, these three sections provide the integrated minor-protection treatment under Articles 28, 34(1)(d), 34(2) and 35 DSA.

EU minor user population

As recorded in the Snapchat Community sub-section of Section 2, Snapchat had an average of [redacted] monthly active recipients aged 13-17 in the European Union during the 2026 assessment period, representing approximately [redacted] of its total EU monthly active recipient population. The minor-user population on Snapchat's in-scope services is therefore both substantial in absolute terms and a discrete protected population to which the elevated mitigation framework recorded in this Section and in Section 5.8 applies.

Likelihood

Underlying risk-category likelihood determinations

The likelihood of negative effects on minors derives from the underlying substantive risk-category determinations across Category 1 and Category 4. The 2026 likelihood, severity and risk-prioritisation determinations across the principal underlying risks are summarised below:

Underlying risk	2026 likelihood
Section 4.1.1 - CSEA Content	Lowest
Section 4.1.2 - Hate Speech	Lowest
Section 4.1.3 - Sale of Prohibited Products - Drugs	Lowest
Section 4.1.3 - Sale of Prohibited Products - Weapons	Lowest
Section 4.1.3 - Sale of Prohibited Products - Other	Lowest

Section 4.1.4 - Dissemination of Terrorist Content	Lowest
Section 4.1.5 - Dissemination of IP-Infringing Content	Lowest
Section 4.1.6 - Adult Sexual Content - Sexual Crimes	Lowest
Section 4.1.6 - Adult Sexual Content - Other	Lowest
Section 4.1.7 - Harassment and Bullying	Lowest
Section 4.1.8 - Self-Harm and Suicide	Lowest
Section 4.1.9 - Violent or Dangerous Behaviour	Lowest
Section 4.1.10 - Harmful False Information	Lowest
Section 4.1.11 - Fraud and Spam	Lowest
Section 4.1.12 - Other Illegal Activities	Lowest
Section 4.4.1 - Public Health	Lowest
Section 4.4.2 - Gender-Based Violence	Lowest
Section 4.4.4 - Physical and Mental Well-Being	Lowest

All underlying risks record their likelihood at the lowest likelihood category. The integrated likelihood of negative effects on minors is therefore driven by the lowest-likelihood determination across the underlying prevalence figures.,

U13 access and use

External Member State evidence. These studies use different age cohorts, survey methodologies, denominators and definitions of platform use and are therefore not directly comparable. They are treated as evidence that access by children below or around the minimum-age threshold is reasonably foreseeable, rather than as estimates of the prevalence of U13 accounts on Snapchat's in-scope services. Publicly available research provides evidence that

[redacted]

access to and use of Snapchat by children below or around the minimum-age threshold is a reasonably foreseeable risk pathway. In Denmark, research published by the Danish Competition and Consumer Authority reported that more than 80% of 12-year-olds had used Snapchat within the preceding 30 days[93]. In Finland, a 2025 KAVI/Verian pilot survey of sixth-grade pupils aged 12–13 reported Snapchat use by 75% of respondents; the report notes that 59% of Finnish children born in 2012 were aged 12 at the time of response[94]. In France, the 2024 Born Social study of 200 children aged 11–12 reported that 36.5% regularly used Snapchat[95]. In Sweden[96]Internetstiftelsen's 2025 research reported that 52% of children in grades 4–6 had used Snapchat during the preceding 12 months, including 40% who reported daily use.

[redacted]

[redacted]

[redacted]

[93] Konkurrence- og Forbrugerstyrelsen (Danish Competition and Consumer Authority), Børn, unge og forældre på sociale medier: Tid, vaner og oplevet forbrug (2 December 2025), pp. 3–4, Figure 2 (reporting that more than 80% of 12-year-olds use Snapchat; estimates are based on reported use of the relevant service within the preceding 30 days).
[94]Kansallinen audiovisuaalinen instituutti (KAVI) / Verian, Lasten ja nuorten mediakysely 2025 (17 June 2025), pp. 2, 9 (pilot survey of sixth-grade pupils aged 12–13; 59% of Finnish children born in 2012 were aged 12 at the time of response; 75% of respondents reported using Snapchat).
[95]heaven x Xerfi-IDM, Born Social 2024 (September 2024), "Game of Platforms" (online survey of n=200 children aged 11–12; 36.5% reported regularly using Snapchat).
[96]Internetstiftelsen, Svenskarna och internet 2025, Chapter 3, "Sociala medier", Diagram 3.30b (2025) (among children in mellanstadiet, 52% had used Snapchat at least once during the preceding 12 months, 46% at least weekly and 40% daily).

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

On that basis, the pathway is assessed within the **lowest likelihood category**, while remaining subject to continued monitoring through the Section 6 risk-detection processes.

The potential consequences for an affected minor may include involvement in drug distribution, organised criminal activity, exploitation or serious violence. The severity of the pathway is therefore assessed as **severe harm industry-wide**. As with the U13 pathway, non-coercive criminal recruitment is not assigned a separate top-level Level 1–3 risk; it forms part of the integrated Negative Effects on Minors assessment and contributes to the existing **Level 1*** overall prioritisation for Section 4.4.3.

5C Typology Cross-reference

Consistent with paragraph 95 and the annexes to the Article 28 Guidelines, which use the 5Cs typology of online risks to children comprising content, conduct, contact, consumer and cross-cutting risks, Snap has applied the 5Cs as a structuring framework for assessing the proportionality of minor-protection measures, including default settings. The table below maps Snap's substantive Article 34(1) risk categories against the 5Cs classification and identifies where each risk pathway is assessed in this Report.

5Cs category	Corresponding Snap Article 34(1) risk categories	Section
Content	CSAM; Illegal Hate Speech; exposure to Prohibited Products or Services; Adult Sexual Content; Self-Harm and Suicide content; Violent or Dangerous Behaviour; Harmful False Information	Sections 4.1.1-4.1.3, 4.1.6 and 4.1.8-4.1.10
Conduct	Harassment and Bullying; self-harm or suicide-related conduct; participation in violent or dangerous behaviour, including harmful challenges	4.1.7-4.1.9
Contact	Grooming and sexual exploitation; contact-initiated CSEA; sextortion and financial sextortion; unwanted or abusive contact with minors; solicitation or recruitment of minors for criminal purposes	Sections 4.1.1, 4.1.6, 4.1.7, 4.4.2 and 4.4.3; related safeguards in Sections 5.4 and 5.8
Consumer	Exposure to or transactions involving Prohibited Products or Services; Fraud and Spam; consumer-protection risks; advertising and other commercial-practices risks affecting minors	Sections 4.1.3, 4.1.11 and 4.2.7; advertising safeguards in Section 5.7
Cross-cutting	Privacy and data-protection risks; geolocation; age assurance and inaccurate age signals; default settings and product design; recommender-system and AI risks; physical and mental well-being	Sections 4.2.3, 4.2.4, 4.2.6, 4.4.3 and 4.4.4; related safeguards in Sections 5.1, 5.6 and 5.8

This mapping operationalises the 5Cs framework within the Report by linking each 5Cs risk pathway to the substantive risk assessment and mitigation sections in which its likelihood, severity and proportionality are assessed.

Minor-specific Violative View Rate (H1 2026)

Minor-specific VVR figures are tracked across the substantive Category 1 categories and are recorded in detail in the relevant underlying Sections and in Section 6.4. The H1 2026 results continue to show very low minor exposure across Snapchat's in-scope services. Across the principal categories for which directly comparable figures are recorded, including Adult Sexual Content, Harassment and Bullying, and Violent or Dangerous Behaviour, the VVR for EU minor users aged 13-17 was lower than the corresponding all-user VVR across the assessed surfaces. The H1 2026 figures therefore do not indicate elevated minor exposure relative to the general user population.

External evidence

The external research base on platform-level minor harm is well-developed and contains both findings of harm and complicating factors. Key sources informing include:

- **APA Health Advisory on Social Media Use in Adolescence (May 2023).** The American Psychological Association concluded that digital platforms are not inherently harmful to young people but that adolescents' experiences online are affected by the features of the platforms they use; the advisory frames the issue as one requiring platform-feature-level mitigation rather than blanket restriction.[97]
- **Committee on the Rights of the Child General Comment No. 25 (2021).** The Committee affirms that the UNCRC applies to online and digital media and identifies that children and teens "need to be protected from risks of all forms of violence in the digital environment" - including "physical or mental violence, injury or abuse, neglect or maltreatment, exploitation and abuse, including sexual exploitation and abuse, child trafficking, gender-based violence, cyber aggression, cyberattacks and information warfare."[98]
- **Council of Europe Recommendation CM/Rec(2018)7.** Identifies the Committee of Ministers' aim of offering "an open, inclusive and secure internet to children and Teens, while ensuring the protection of their human rights."[99]
- **Ofcom Children Register of Risks.** Documents children's exposure to abuse and hate content and the associated service-level risk factors. Ofcom reports that, among respondents aged 13 and over, 78% of those whose most recent experience involved hateful, offensive or discriminatory content, and 69% of those whose most recent

[97] American Psychological Association, Health Advisory on Social Media Use in Adolescence (May 2023).
[98] UN Committee on the Rights of the Child, General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment, UN Doc CRC/C/GC/25 (2 March 2021), paras 80-82.
[99] Council of Europe, Committee of Ministers, Recommendation CM/Rec(2018)7, Guidelines to Respect, Protect and Fulfil the Rights of the Child in the Digital Environment (adopted 4 July 2018).

[redacted]

experience involved misogynistic content, said that the experience occurred on a social-media website or app. Ofcom notes that these samples included both adults and children because the child-only base was too low to report separately.[100]

- **LSE EU Kids Online (February 2025).** Research on the children's-rights implications of EU social-media restrictions, identifying that overly broad age-based restrictions can themselves engage children's rights to expression, information and participation.[101]
- **Eric Goldman, "Segregate-and-Suppress" (SSRN).** Identifies that overly restrictive minor-targeted speech regulations may infringe minors' fundamental rights to expression and access to information.[102]
- **ThinkYoung research (February 2025).** Research on European youth perspectives on online platforms, including direct evidence of teenager perspectives on platform participation and its role in their development and social connection.[103]
- **Snapchat-specific research, University of Amsterdam.** A 100-day diary study involving 479 adolescents found that effects differed materially by platform, person and outcome. Its exploratory platform-specific analyses identified positive or null effects associated with Snapchat across well-being, self-esteem and friendship closeness, supporting the need to distinguish between different platform designs rather than treating social-media use as a uniform experience.[104]
- **Research concerning illicit-drug promotion and access through social media.** WIRED, EMCDDA and Europol, and peer-reviewed research document the growing use of social-media and instant-messaging services for illicit-drug advertising, sales and distribution, including the particular risk that mainstream platforms facilitate exposure and access among children and adolescents.[105]

Member State age-assurance framework engagement

Snap continued to assess evolving age-assurance requirements and regulatory expectations at EU and Member State level, including the European Commission's Article 28 Guidelines and relevant national frameworks. [redacted]

[100] Ofcom, Children's Register of Risks (24 April 2025), paras 5.77-5.79 and nn 689-692.
[101] EU Kids Online, Protecting, Not Excluding: Why Banning Children from Social Media Undermines Their Rights (February 2025).
[102] Eric Goldman, 'The "Segregate-and-Suppress" Approach to Regulating Child Safety Online' (2025) 28 Stanford Technology Law Review 173, DOI: 10.2139/ssrn.5208739.
[103] Sameer Hinduja and Farah Lalani, Empowering and Protecting European Youth Online: Streamlining Legislation and Promoting Positive Digital Experiences (February 2025), supported by ThinkYoung; ThinkYoung, Usage of Online Platforms & Youth Regulation: European Teen & Youth Focus Group Report (2025).
[104] Amber van der Wal, Ine Beyens, Loes H.C. Janssen and Patti M. Valkenburg, 'Diverse Platforms, Diverse Effects: Evidence from a 100-Day Study on Social Media and Adolescent Mental Health' (2026) 45 Current Psychology, article 36, DOI: 10.1007/s12144-025-08893-7.
[105] Mattha Busby, 'Drug Dealers Have Moved on to Social Media', WIRED (17 December 2024); European Monitoring Center for Drugs and Drug Addiction and Europol, EU Drug Market: Drivers and Facilitators (30 May 2024), 'Technology and Innovation', doi:10.2810/219567; Ashly Fuller, Marie Vasek, Enrico Mariconti and Shane D Johnson, 'Understanding and Preventing the Advertisement and Sale of Illicit Drugs to Young People Through Social Media: A Multidisciplinary Scoping Review' (2024) 43(1) Drug and Alcohol Review 56-74, doi:10.1111/dar.13716.

[redacted]

Likelihood conclusion

On the basis of the underlying Category 1 and Category 4 likelihood determinations (lowest across all underlying risks), the very low minor-specific VVR figures, the absence of material reports identifying minor-specific harms as elevated risk on Snapchat's in-scope services, and the structural minor-protection framework recorded in Section 5.8, the likelihood of negative effects on minors on Snapchat's in-scope services is assessed at the **lowest likelihood** category. There is no change in this determination from the 2025 Report. The 2026 Risk Assessment preserves the recognition that the integrated minor-protection framework operates across an architecture (rather than at any single point) and that no single likelihood metric captures the full systemic-risk profile.

Severity

The severity of negative effects on minors warrants elevated assessment relative to the same risks affecting adult users, on three analytical bases:

- **Vulnerability of the population.** Minors and teens (13-17) are recognised as a vulnerable population by virtue of their developmental stage, evolving cognitive and emotional capacities, and limited capacity to assess and respond to certain online risks. Where harm materialises, the impact on a minor can be more severe than the equivalent harm to an adult, and can have lasting effects on development, education, and digital confidence.
- **Categorical recognition by international human rights frameworks.** Article 24 of the EU Charter, the UNCRC and General Comment No. 25 (2021), the Council of Europe Recommendation CM/Rec(2018)7, and Member State Age Appropriate Design Codes all categorically recognise the elevated standard of care owed to minors in the digital environment.
- **Categorical recognition by the DSA.** Article 28 DSA imposes a specific obligation on platforms accessible to minors; Article 34(1)(d) lists protection of minors as a discrete systemic-risk category; Recital 71 DSA identifies minors as a particularly affected category requiring elevated attention.

[106] Snap Inc., '[Launch doc] [EU] Default selected date to today with 24h cooldown on Birthday page During Reg' (22 June 2026); Snap Inc., 'Study Launch Document - REGISTRATION_AGE_GATING_V2' (study conducted 28 May-7 June 2026; EU iOS rollout scheduled for 22 June 2026).

[redacted]

On these analytical bases, severity for negative effects on minors is elevated by reference to each underlying substantive category. The elevated-severity application is recorded in detail below:

Underlying risk category	Industry-wide severity (general population)	Elevated severity for minors
CSEA Content (4.1.1)	Severe	Yes - by definition affects minors; severe in baseline classification.
Hate Speech (4.1.2)	Significant	Yes - minors are more impressionable and may experience identity-development harm if exposed.
Sale of Prohibited Products (4.1.3)	Severe (drugs) / Serious (weapons) / Significant (other)	Yes - minors are a documented target population for online drug-supply pathways.
Adult Sexual Content (4.1.6)	Serious (sexual crimes) / Significant (other)	Yes - minors experience particular harm from exposure to sexual content; sextortion specifically targets minors.
Harassment and Bullying (4.1.7)	Significant	Yes - cyberbullying impact on minor mental health and educational outcomes is well-documented.
Self-Harm and Suicide (4.1.8)	Serious	Yes - elevated to severe for minors; the principal vulnerable population for self-harm content exposure.
Violent or Dangerous Behaviour (4.1.9)	Significant	Yes - viral-challenge content disproportionately affects minor populations through imitation pathways.
Harmful False Information (4.1.10)	Significant	Yes - minors are still developing critical-evaluation capacities and may be more susceptible to misinformation harm.
Public Health (4.4.1)	Severe	Yes - minors are particularly affected by health misinformation, including in connection with substance-use prevention.
Gender-Based Violence (4.4.2)	Serious	Yes - financial sextortion of teenage boys identified as a documented organised-crime targeting pattern.
Physical and Mental Well-Being (4.4.4)	Severe	Yes - minors are the principal population engaged by the well-being framework.

[redacted]

Recognising the across-the-board elevated-severity multiplier and the categorical recognition by international and EU frameworks, the severity of negative effects on minors is assessed as **severe harm industry-wide**, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems take user age into account to provide age-appropriate recommendations. [redacted] The Content Guidelines for Recommendation Eligibility apply additional standards beyond the Community Guidelines, calibrated to a child-population baseline, with the operational assumption that users as young as 13 may be viewing content. Further detail is provided in Section 5.6. Snap provides a Reset Suggested Content feature through which users, including minor users, can refresh their Spotlight and Discover recommendations by resetting most signals used to personalise public-content feeds. [redacted] EU users, including minor users, may also disable personalised recommendations for Spotlight and Discover through European Union Controls. Where personalisation is disabled, recommendations are based on basic information such as language, age and country rather than behavioural engagement signals. Snap has therefore implemented both a feed-reset mechanism and a non-personalised feed option. The non-personalised option is available for users to opt in to and is not the default setting for any users.

[redacted]

Content moderation systems	Snap operates specific proactive and reactive moderation procedures relevant to minors. These include age-gating of sensitive content; escalation guidance for moderators on minor-related cases; the dedicated “It involves a child” reporting reason; the Family Center settings allowing parents/guardians to adjust access to sensitive material and review their teen’s activity; and the expansion of profile-level moderation (Section 5.4.4) supporting detection of patterns engaging minor users at the username, display-name and profile level. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Snap Terms of Service require public content to be suitable for all Snapchatters (i.e. 13+). The Community Guidelines prohibit account creation by users under 13 and apply to all users equally. Teen-friendly Terms language and support pages are maintained to support comprehension by users in the 13-17 cohort. [redacted] Further detail in Section 5.2 and Section 5.5. Enforcement of these standards reduces the risk of minors being exposed to age-inappropriate content by removing or restricting content that violates the Community Guidelines and by applying stricter recommendation-eligibility, age-gating and filtering controls to public content shown to minor users. The effect of those controls is considered alongside the minor-specific VVR and other effectiveness evidence recorded in this Section and in Sections 5.4, 5.6 and 5.8.
Systems for selecting and presenting advertisements	Specific minor-protection advertising measures include: the prohibition on advertising-targeting based on minor users’ activity data; [redacted] the discontinuation of EU political advertising as of October 2025; and [redacted]. Further detail in Section 5.7.

[redacted]

Data-related practices	PASS Review and DDIA framework apply enhanced minor-protection scrutiny to product and feature changes affecting minor users. [redacted] The prohibition on processing minor users' activity data for advertising-targeting purposes is the principal data-related minor-protection commitment. Further detail in Section 4.2.4 and Section 6.3.
Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of minor-targeted or minor-affecting content is structurally constrained by the platform architecture (recommended content in public feed is moderated; no broad reshare functionality of Stories; pre-distribution moderation for content reaching wide audiences; content flagged as sexually suggestive or otherwise sensitive is filtered from Teen feeds, the bidirectional friending requirement; the contact-restriction defaults for minor accounts). Snap has also implemented specific proactive and reactive moderation procedures to prevent minors from accessing content that is harmful to them. Content recommended to users on Spotlight and Discover, our video sharing platforms, is moderated using a combination of auto-moderation and human moderation, and is human moderated before being widely distributed. Lens and Ads are subject to review processes.

[redacted]

	The minor-specific Violative View Rate figures recorded in the Likelihood discussion above show very low minor exposure to violating content across surfaces and substantive categories.
Regional and linguistic aspects	Regional differences are addressed through jurisdiction-specific minimum-age requirements and market-specific implementation [redacted]

Overall Potential Risk Prioritization

Applying the framework in Section 3.5, the default matrix output for severe harm industry-wide combined with lowest likelihood is Level 3. However, applying the case-by-case adjustment methodology described there, the risk of negative effects on minors is assessed at Level 1 overall potential risk prioritisation. The upward adjustment reflects the particular vulnerability of the minor-user population; the potential for severe and lasting effects on development, physical and mental well-being and safety where harms materialise; the specific recognition of minors as a protected population under Article 24 of the EU Charter and Articles 28 and 34(1)(d) DSA; and the integrated nature of the content, conduct, contact, consumer and cross-cutting risk pathways. The adjustment therefore reflects the potential gravity and breadth of harm to minors rather than an elevated observed prevalence on Snapchat. There is no change in this determination from the 2025 Report. The operational refinements reflect the continuous evolution of Snap's minor-protection programme rather than a change in the underlying risk determination.

[redacted]

[redacted]

[redacted]

Snap's Mitigations

Three-pillar minor-protection framework

Snap's minor-protection mitigation framework operates around three core pillars:

[redacted]

Age-Appropriate Design Code adherence

Snap's practices align with principles reflected in age-appropriate design frameworks, including the United Kingdom's Children's Code and relevant European frameworks, and Snap applies those principles as a baseline across its products and services.

- Act in the best interests of teens;
- Respect teens' freedom to express themselves safely;
- Recognise teens' right to information about the world;
- Position parents/guardians and trusted adults to help guide teens in the responsible exercise of their rights on the platform;
- Attach a heightened safety interest to teens using the products, with unique safeguards for young users;
- Establish operational processes to ensure that products are developed in a way that upholds these tenets.

In addition to these AADC tenets, Snap consistently enforces its policies disallowing users under the age of 13 from creating or maintaining an account. Persons younger than 13 are blocked from

[redacted]

[redacted]

Snapchat is designed with Teens safety and privacy in mind

Snap's approach to privacy and safety by design means that we generally design for our youngest users first and work upwards from there. This means our first layer of protection for minors are the mitigations that are designed for Teens but apply to all to advance safety across our community. These include:

Terms

- This helps all users to understand what activity is prohibited on Snapchat and the consequences, which reduces the likelihood they will engage in illegal or violating activity.
- Our Term of Service requires public content to be suitable for all Snapchatters (i.e. 13+). This helps ensure Snapchat is suitable for Teens and non-Teen users and reduces risks across all of the risk categories identified for Snapchat.

Transparency

- Our privacy and safety center and support pages as explained in the Transparency part of this Report have also been designed to be read by all users, from 13 years upwards. This helps all users to understand how Snapchat works, what options they may have, how we moderate and enforce our terms and how they can get support. This reduces both the likelihood of illegal or violating activity and the severity of harm in the event they are exposed to illegal or violating activity despite our limitations.
- For example, Teens reading our privacy center can understand how their data is being processed by Snap and find more information about relevant privacy settings which reduces the likelihood and severity of negative effects on Teens' data protection rights.

Moderation and Enforcement

- We have dedicated processes in place to moderate content and enforce our Terms, as explained earlier in this Report. These help prevent negative effects across all of the risks we have identified for Teens, as well as other Snapchatters, using Snapchat.

Reporting and Blocking:

- Teens have the ability to report abuse they may observe or experience within Snapchat. [redacted]

Product design and functionality:

- Snapchat has been designed and adapted to limit risks to Teens and other Snapchatters.

[redacted]

Teen specific mitigations

In addition to the mitigations that have been designed for Teens but apply to all, we also have additional protections that apply just to our Teen users to help mitigate increased risks for certain features or issues. These include:

Terms and Transparency:

- We recognise the need for Teens to understand the functioning of the platform. Our Privacy Center was designed for our youngest Snapchatters and was intentionally developed to be easy to read and understood by all members of our community. We created our privacy and safety hub, with pages such as our Privacy by Product page, to give Snapchatters a high-level summary of our privacy and safety practices across each of our products and features. We also created a video to visualize our privacy practices, and use icons and other best practices as recommended by privacy and safety experts and the recognised Age Appropriate Design Codes.

Public Profiles:

- Older teenagers (ages 16–17) in the EU, EEA, and UK have access to Public Profile and public posting options. This means they can create or update a publicly accessible profile, share Public Stories, and submit content for Spotlight recommendations.
- [redacted]

Publishing Public Content:

- Anyone who shares content on Spotlight must agree to Snap Spotlight Submission and Revenue Terms, which also requires collection of appropriate consent from anyone appearing in the Snap. We have put in place additional protections and limitations on Spotlight. [redacted]

[redacted]

- Users can disable comments on any post. [redacted]. When Teens post to Spotlight or Snap Map, their profile details are anonymized as an extra precaution. This reduces the likelihood of a range of risks identified and applicable to Teens - for example Harassment and Bullying derived from public profiles, risk of Negative Effects on Physical and Mental Health due to pressures to post, to inappropriately contacted by malicious adults seeking to engage in activity prohibited on Snapchat.

*Accessing Public Content**:***

- Teen accounts are restricted from access to certain content that is generally considered suitable for 13+ but may contain certain shocking or sensitive content some may not find appropriate. For example, we moderate Spotlight content and prevent sexually suggestive content from appearing to Teens. [redacted].
- On Discover, our policies are written with the understanding that people as young as 13 may be viewing the content and should be age appropriate. If it isn't, our partners have tools to make their content 18+ at their discretion. [redacted]
- These steps reduce the likelihood of Teens accessing illegal or violating content, or content that may negatively affect their rights, security and health.

Maps:

- [redacted]

Lenses:

- Snap takes steps to help Teens understand the design and functioning of Lenses and the risks inherent with Lenses that may cover controversial topics (i.e., Beautification Lenses). Snap rejects Lenses that attempt to realistically imitate cosmetic surgery and sponsored

[redacted]

Lenses for cosmetic surgery. On the other hand, Snap allows cartoon-like Lenses (big eyes, little face) on the platform because of their whimsical and non-photo-realistic nature.

- Additionally, Snap uses context cards attached to sending or posting Snaps so the receiver knows a Lens has been applied. For example, this has been done in relation to the 'Baby Lens'. This Lens can turn an adult user into looking like a baby. There were concerns that adults could use this Lens to pretend to be younger and connect with Teens to "groom" them. To prevent this from happening, a rattle was added to the Lens creative to clearly indicate that it is a Lens which has transformed the person into a baby.
- Also, the context card indicates to on-platform viewers that a Lens has been applied to the Snap. The context card attaches to sending or posting Snaps so the receiver knows a Lens has been applied.
- In addition, Snap geofences or age-gates certain Lenses (e.g. related to alcohol, gambling, NFTs, etc.). For example, a Vivino Lens will only be shown to 18+ users in the EU, mitigating the risk of such content being shown to users that are under the legal drinking age.
- Snap also respects Teens' right to privacy in the context of Lenses, for Lenses does not require facial identification (object detection is sufficient) and data collected from Lenses is only used to provide the service. This reduces the likelihood and severity of negative effects on Teen's data protection rights.
- Respecting Teen's right to non-discrimination, Snap designs every Lens with race, gender, ethnicity and cultural norms in mind. Snap leverages its ever-growing diverse training datasets, as well as feedback from community members. If a Lens does not resonate with our community, as expressed through a high ratio of user reports, we take that feedback into consideration and will re-review the Lens with a goal to leave as-is, modify, or remove.
- Snap moderates all Lenses available on the platform before availability [redacted] This reduces the likelihood of Teens accessing illegal or violating content, or content that may negatively affect their rights, security and health.

Advertising:

We take additional measures to mitigate risks raised by Snapchat Ads in the following ways:

- Advertisers must comply with our Ad Terms, Advertising Policies, and applicable national advertising codes. We prohibit ads that address or intend to appeal specifically to children under the age of 13.
- We review every ad before it is published on Snapchat to ensure compliance with the Policy. We provide moderators with Global Ad Review Content Policies (which are country-by-country policies for reviewing appropriate ads).
- Our review takes into account the targeted audience [redacted]

[redacted]

Family Center:

We recognise that parents play a critical role in shaping their children's use of technology.

- Through Family Center, Snap provides a way for parents/guardians/trusted adults to gain insight into their teen's in-app friends, communications and certain other activities on Snapchat, creating opportunities for meaningful discussion between teens and adults about responsible decision-making online.
- Family Center allows parents/guardians/trusted adults to, among other things: (i) view their teen's existing friends list and new friends that were added; (ii) see which friends their teen has recently communicated with; (iii) see key privacy settings like birthdate, Public Profile, and location sharing settings; (iv) more easily report concerns and (v) turn off communication with My AI; and (vi) request location sharing.

Article 35(1) DSA mitigation mapping

The mitigations in place for this risk are mapped to Article 35(1)(a)-(k) DSA below. The substantive minor-protection framework is recorded in detail in Section 5.8 (Protection of Minors); the mitigation table provides the Article 35-mapping view of the framework, with cross-references to Section 5.8. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. The architectural design of Snapchat reflects minor-protection considerations as foundational: bidirectional friending requirement; contact-restriction defaults for minor accounts; the absence of an open recommender feed for unvetted user-generated content; private friend lists by default; privacy-by-default Snap Map settings for minor accounts; minor-specific JIT notices warning of contact from accounts in elevated-scamming regions. The minor-account default settings (Section 5.1.3), onboarding redesign leading with safety features, and friending-protection refinements, including the reduction of the Find Friends candidate strike-count threshold from N=3 to N=2 (effective May 2026), strengthen the design framework.
Terms and Enforcement - Art. 35(1)(b)	Applies. The Snap Terms of Service require public content to be 13+ suitable; the Community Guidelines prohibit accounts by users under 13 and apply uniformly to all users; teen-friendly Terms language and support pages support comprehension by 13-17 users; and [redacted]. Further detail in Section 5.2 and Section 5.5.

[redacted]

Moderation - Art. 35(1)(c)	Applies. Specific proactive and reactive moderation procedures for minor protection include: age-gating of sensitive content; the dedicated "It involves a child" reporting reason; escalation guidance for moderators on minor-related cases; [redacted] Further detail in Section 5.4 and Section 4.1.1.
Algorithmic Systems - Art. 35(1)(d)	Applies. Recommender systems apply the Content Guidelines for Recommendation Eligibility, calibrated to a child-population baseline. Spotlight and Discover content for minor users is subject to stricter eligibility rules and content filtering. Algorithmic systems do not categorise users on the basis of special category data (Article 9 GDPR) for advertising or recommendation purposes. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Specific minor-protection advertising measures: prohibition on advertising-targeting based on minor users' activity data; [redacted] discontinuation of EU political advertising as of October 2025; pre-publication review of all advertising with category-based filtering for minor exposure. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. The Section 6 ongoing risk-detection processes - including the Safety Advisory Board (which includes child-safety experts); engagement with civil-society organisations focused on child protection; the PASS Review and DDIA framework with enhanced minor-protection scrutiny; minor-specific Violative View Rate tracking across substantive categories; and Prevalence Testing across minor-affecting categories - monitor risks bearing on this category. Further detail in Section 6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA may concern substantive harms affecting minors and are processed through the applicable priority-notice procedures. Snap also cooperates with specialist child-safety organisations, including [redacted] through the separate operational frameworks described in Sections 4.1.1 and

[redacted]

	5.10. These mechanisms support the broader minor-protection framework. Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap engages with the European Commission and Member States on Article 28 Guidelines implementation; participates in the EU Internet Forum (EUIF); engages with civil-society organisations focused on children's digital rights [redacted]); follows relevant age-appropriate design frameworks, including the United Kingdom's Children's Code and European frameworks; in 2017 Snap joined FSM (the German Voluntary Self-Regulation of Multimedia Service Providers) and signed the FSM Code of Conduct. The Content Crisis Response Protocol (Section 5.12) operates across categories engaging minor users. Further detail in Section 5.12.
Transparency - Art. 35(1)(i)	Applies. Information for minor users and their parents/guardians is provided through: the Teen-friendly Community Guidelines and Terms summary; the Family Center support pages and the parents site; the Safety Center and Privacy Center; Just-In-Time (JIT) notices presented during onboarding and at sensitive product moments; in-app educational resources including Here for You (mental health and well-being support); and the Heads Up tool for drug-related search interventions. Snap's Transparency Reports include minor-specific data where relevant (in particular, minor-user Violative View Rate figures across substantive harm categories). Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. The full minor-protection framework is recorded in Section 5.8 (Protection of Minors), structured as an integrated multi-layer architecture: app-store level safeguards; device-level safeguards (operating system parental controls and family tools); platform-level safeguards (account-creation under-13 prevention; [redacted] contact-restriction defaults; bidirectional friending); product-level safeguards (Lenses, Spotlight, Discover, Public Profiles, Snap Map, advertising); reporting and blocking; private messaging and friending controls; Family Center and parent tools, [redacted]; and the Article 28 Guidelines mapping. Further detail in Section 5.8.

[redacted]

Content Authenticity - Art. 35(1)(k)	Applies. Generative-AI features made available to minor users are subject to additional content-authenticity safeguards: input/output scanning; retry-logic for unsafe outputs; JIT notices warning of AI-generated character of outputs; watermarking of generative-AI outputs (with the icon displayed on Lenses that manipulate an image of a Snapchatter to make them look younger). Snap has taken steps to mitigate the risk that (i) its generative-AI tools are used for creating illegal or otherwise minor-affecting content, and (ii) such content created using generative-AI tools on any online platform is disseminated on Snapchat's in-scope services. Further detail in Section 5.9.

Article 28 DSA Guidelines (Commission Guidance, 2025)

Following the publication by the European Commission of its Guidance on Article 28 DSA in 2025, Snap has assessed Snapchat's in-scope services against the Guidance's principal recommendations. The detailed Article 28 Guidelines mapping is recorded in Section 5.8, where Snap's measures are mapped against the eight principal sections of the Guidance: General Principles; Risk Review; Age Assurance; Defaults and Account Settings; Interfaces, Tools and AI; Recommender Systems and Content Moderation; Commercial Practices; Reporting Support and Guardian Tools.

In summary, Snap's existing measures align with the Guidance recommendations across the principal sections. Where specific recommendations are not currently implemented, Snap has either: (i) determined that existing measures are proportionate and effective in light of the specific nature of Snapchat's in-scope services, or (ii) recorded the recommendation as an item under continued consideration. [redacted]

These refinements operate together to strengthen the integrated minor-protection framework recorded in Section 5.8 across the broader architecture. The framework's analytical determinations (likelihood, severity, risk prioritisation) are unchanged from the 2025 Report.

Conclusion

[redacted]

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above, Snap has concluded that the mitigations in place for the protection of minors on Snapchat's in-scope services are reasonable, proportionate and effective.

[redacted]

There is no change in the overall conclusion from the 2025 Report. The operational refinements strengthen the framework, while preserving the proportionality and the architectural integrity of the existing minor-protection framework. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to: the operational outputs of the minor-protection refinements; the implementation framework of the Article 28 Guidelines e; the academic and civil-society research base on the appropriate balance between protection and participation in respect of children's digital lives; the Member State age-assurance frameworks under development; and the continued evolution of generative-AI-related minor-protection risk patterns.

Snap continues to support the Commission's broader minor-protection work and the development of the Article 28 Guidelines implementation framework. This engagement will continue through subsequent assessment periods.

4.4.4 Serious Negative Consequences on Physical and Mental Well-Being

This Section 4.4.4 sets out Snap's assessment of the actual or foreseeable serious negative consequences on the physical and mental well-being of recipients of the service arising from the operation of Snapchat's in-scope services. Article 34(1)(d) DSA identifies serious negative consequences on physical and mental well-being, alongside negative effects on public health,

[redacted]

gender-based violence, and protection of minors, as a Category 4 systemic-risk category subject to assessment and mitigation.

This risk is closely related to, and should be read together with, Section 4.4.1 (Public Health), Section 4.4.3 (Minors), and the underlying Category 1 sections that relate to wellbeing-relevant conduct, in particular, Section 4.1.6 (Adult Sexual Content), Section 4.1.7 (Harassment and Bullying), Section 4.1.8 (Self-Harm, Suicide and Eating Disorders), Section 4.1.9 (Violent or Dangerous Behaviour), and Section 4.1.10 (Harmful False Information). If we follow our matrix, we would qualify the overall potential of this risk as Level 3, however, given the importance of this issue, especially in relation to younger users, we have decided to deviate from the matrix, and marked the potential risk as a **Level 1 priority**. There is no change in this assessment from our 2025 Report.

Likelihood

The likelihood determination for this risk draws on three principal evidence bases: (i) underlying Category 1 likelihood determinations and operational data; (ii) the contemporary external research base on platform-wellbeing relationships; and (iii) Snap's own Digital Well-Being Index research.

Underlying Category 1 likelihood determinations

The substantive risk pathways for this risk are addressed across multiple underlying Category 1 risks. The 2026 likelihood determinations across these underlying risks are uniformly recorded at lowest likelihood: Section 4.1.6 (Adult Sexual Content), Section 4.1.7 (Harassment and Bullying), Section 4.1.8 (Self-Harm, Suicide and Eating Disorders), Section 4.1.9 (Violent or Dangerous Behaviour), and Section 4.1.10 (Harmful False Information). The empirical observations recorded across these risks, sustained low PVP figures, very low VVR figures (including for EU minor users), rapid median turnaround times for enforcement action, and the operational performance of the Proactive Safety Detection System, support the determination that the volume of content engaging wellbeing harm on Snapchat's in-scope services is low.

Self-Harm and Suicide-specific PVP was [redacted] for the 90-day period ending 30 June 2026, compared with [redacted] recorded in the 2025 Report. As noted in Section 6.4, the two periods are not fully like-for-like because Snap re-baselined its prevalence sampling in March 2025. The 2026 result nevertheless remains low in absolute terms, and the corresponding VVR results for EU users and minors were also very low.

External research

The research base on platform-wellbeing relationships is genuinely contested and has continued to evolve materially during the 2025-2026 period. This section presents the research base evenhandedly, recognising both the harms identified across multiple research and policy reviews

and the methodological complications that have prevented the field from reaching settled conclusions.

Authoritative research and policy reviews. Several authoritative bodies have published comprehensive reviews of the platform-wellbeing research base during the 2023-2025 period:

- The **US National Academies of Sciences, Engineering, and Medicine (2024)** review of social media use and adolescent mental health concluded that the available evidence does not support a determination that social media use causes adolescent mental health problems on a population basis, while identifying that specific subgroups and use patterns can present elevated risk;[107]
- The **Norwegian Screen Time Committee (Skjermtidsutvalget, 2024)** review for the Norwegian Government found that the available evidence on the relationship between screen time and adolescent wellbeing is mixed and that policy responses should be calibrated to specific use patterns rather than generic time-based restrictions;[108]
- The **Belgian Superior Health Council (Conseil Supérieur de la Santé / Hoge Gezondheidsraad, 2025)** report on screen use among children and adolescents recommended a calibrated policy approach distinguishing between use patterns rather than generic restrictions;[109]
- The **French ANSES (Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, October 2025)** expert collective report "Adolescents et réseaux sociaux" (Reference 2019-SA-0152) provided a comprehensive synthesis of the platform-adolescent wellbeing research base in the French and Member State context, with detailed treatment of measurement methodologies, exposure metrics, and outcome categories. The report identifies methodological complexity in the research base and supports calibrated policy responses;[110]
- Nature Reviews Psychology (2024). A review by Orben et al. found that research linking social media use to adolescent mental health has produced mixed and inconsistent findings, with effects likely varying according to individual vulnerability, developmental stage, platform features and patterns of use. The review supports a mechanism-specific and context-sensitive approach rather than broad conclusions based solely on overall time spent on social media.[111]

[107] National Academies of Sciences, Engineering, and Medicine, Social Media and Adolescent Health (National Academies Press, 2024), doi: 10.17226/27396.
[108] Norwegian Ministry of Education and Research, NOU 2024: 20, Det digitale (i) livet: Balansert oppvekst i skjermenes tid (11 November 2024).
[109] Belgian Superior Health Council, The Effects of the Use of Screens and Social Media on Young People, Opinion No. 9877 (4 December 2025).
[110] ANSES, Avis et rapport sur les usages des réseaux sociaux numériques et santé des adolescents, Reference 2019-SA-0152 (16 December 2025).
[111] Amy Orben, Adrian Meier, Tim Dalgleish et al., "Mechanisms Linking Social Media Use to Adolescent Mental Health Vulnerability," Nature Reviews Psychology 3 (2024): 407–423, doi: 10.1038/s44159-024-00307-y.

[redacted]

- The **BMJ (March 2025)** published a synthesis from international experts identifying methodological limitations in the existing research base and recommending caution in inferring causation from correlation in adolescent screen-use studies.[112]

Snapchat-specific research. Recent academic research includes Snapchat-specific findings that are particularly relevant to this risk:

- **University of Amsterdam (2026).** A 100-day daily-diary study of 479 adolescents found that the effects of social media use differed materially across platforms. In exploratory platform-specific analyses, time spent on Snapchat was associated with positive effects on friendship closeness and well-being and no significant effect on self-esteem. The study also identified negative effects from overall social media use for many participants, so its findings support a differentiated, platform-specific assessment rather than a general conclusion that social media use is beneficial.[113]
- **Australian Future Proofing Study (2024).** Analysis from the Australian Future Proofing Study found that greater time spent on Snapchat was not significantly associated with the mental-health symptoms examined. It also found that more frequent use of social media to communicate with people adolescents knew offline was associated with lower symptoms of depression and anxiety. These findings support distinguishing reciprocal communication with known contacts from other patterns of social media use, but do not establish that Snapchat use caused improved mental-health outcomes.[114]
- **University of Antwerp research on Streaks.** Research involving 2,483 early adolescents found associations between Streak engagement, fear of missing out, problematic smartphone use and social-media self-control, although the reported relationships between these measures and the number or duration of Streaks were generally weak.[115]
- **ThinkYoung (February 2025).** Direct survey research on European youth perspectives on online platforms identified that adolescents themselves perceive platform participation as supportive of social connection and personal development, with particular nuance about how different platforms serve different needs.[116]

[redacted]

[112] Victoria A. Goodyear et al., “Approaches to Children’s Smartphone and Social Media Use Must Go Beyond Bans,” BMJ 388 (27 March 2025): e082569.
[113] Amber van der Wal, Ine Beyens, Loes H. C. Janssen et al., “Diverse Platforms, Diverse Effects: Evidence from a 100-Day Study on Social Media and Adolescent Mental Health,” Current Psychology 45 (2026), Article 36, doi: 10.1007/s12144-025-08893-7.
[114] ReachOut, Beyond Blue and Black Dog Institute, Submission 168 to the Australian Parliament Joint Select Committee on Social Media and Australian Society (2024), pp. 9–10.
[115] Christina M. van Essen and Joris Van Ouytsel, “Snapchat Streaks: How Are These Forms of Gamified Interactions Associated with Problematic Smartphone Use and Fear of Missing Out Among Early Adolescents?”, Telematics and Informatics Reports 11 (2023): 100087.
[116] ThinkYoung, Usage of Online Platforms & Youth Regulation: European Teen & Youth Focus Group Report (2025).

[redacted]

Research on the implications of restrictive policy responses. Recent academic and civil-society research has identified that overly restrictive minor-targeted platform measures may themselves engage wellbeing concerns by infringing children's rights to expression and access to information:

- The **LSE EU Kids Online research (February 2025)** on EU-wide social-media restrictions identified that overly broad age-based restrictions can themselves engage children's rights to expression, information and participation;[117]
- Eric Goldman, **"Segregate-and-Suppress" (SSRN)**, identifies the constitutional and other implications of age-based online speech restrictions, supporting the position that calibrated platform measures (rather than restrictions) are the appropriate policy frame;[118]
- Research by Candice Odgers (**Nature, March 2024**), "Reframing the Social Media Question," supports the analytical position that the policy debate has often outrun the empirical foundations and recommends a calibrated approach to platform-design measures.[119]

Earlier research and policy positions. The broader evidence base considered for this assessment also includes the US Surgeon General's May 2023 Advisory on social media and youth mental health, Hancock et al. (2022), Valkenburg et al. (2022), Orben et al. (2020), and research and commentary published by the Harvard T.H. Chan School of Public Health. These sources provide additional context regarding both potential harms and the methodological limitations affecting conclusions about population-level causation.

External expert and youth input. In addition to this published research, Snap supplements its internal operational and survey evidence with structured external expert and youth input. This includes consultation with Snap's Safety Advisory Board, whose membership includes academics, researchers, safety professionals and other experts across a range of online-safety disciplines, as

[117] EU Kids Online, Protecting, Not Excluding: Why Banning Children from Social Media Undermines Their Rights (February 2025).
[118] Eric Goldman, "The 'Segregate-and-Suppress' Approach to Regulating Child Safety Online," 28 Stanford Technology Law Review 173 (2025).
[119] Candice L. Odgers, "The Great Rewiring: Is Social Media Really Behind an Epidemic of Teenage Mental Illness?", Nature 628 (29 March 2024): 29–30.

well as the Council for Digital Well-Being and Teen Council. Snap also engages with civil-society organisations and academic researchers on youth well-being and platform-wellbeing issues. These external inputs inform the identification and assessment of emerging risks and, where relevant, the design and evaluation of mitigations.

Snap's Digital Well-Being Index (DWBI)

To gain insight into how teens and young adults experience online platforms across the broader digital ecosystem, Snap conducts Digital Well-Being Index survey research across three age groups in six countries. The research provides an additional input for identifying trends in users' reported online experiences and informing Snap's digital-well-being initiatives. No quantitative conclusion is drawn from DWBI measurement outputs for the 2026 assessment period. Further detail on the methodology and findings is provided in Section 6.6.

Likelihood conclusion

On the basis of: (i) the underlying Category 1 likelihood determinations recorded at lowest likelihood across all five principal categories engaging wellbeing harm; (ii) the operational performance of the moderation framework (sustained low PVP figures, very low minor-specific VVR figures, rapid median turnaround times); (iii) the contemporary external research base, which has materially evolved during 2025 toward a more nuanced view of the platform-wellbeing relationship and away from the earliest causal-inference framings; (iv) the Snapchat-specific research identifying positive associations between platform use and adolescent social connection; (v) Snap's Digital Well-Being Index research providing an additional source of information on reported online experiences across the broader digital ecosystem; and (vi) the structural absence on Snapchat of the platform features (open recommender feeds, user-generated content does not automatically advance from one user's content to the next user's content, follower-count gamification) most consistently identified across the research base as drivers of compulsive use and negative wellbeing outcomes, the likelihood of serious negative consequences on physical and mental well-being on Snapchat's in-scope services is assessed at the **lowest likelihood** category. There is no change in this determination from the 2025 Report.

Severity

Snap recognises that negative experiences on digital platforms can have material consequences for mental health and wellness. Although the platform-wellbeing research base is contested in respect of population-level effects, the specific harm pathways within this category are well-documented and can produce serious consequences:

- compulsive use patterns, where engagement-maximising design features can produce dependency-like behavioural patterns, particularly among adolescents and users with pre-existing vulnerabilities;

- body-image and self-esteem effects, where appearance-focused content (including Lens AR features that alter appearance) can produce body-image dissatisfaction and related psychological harm;
- anxiety, depression and sleep-quality effects, identified in research on adolescent populations with elevated platform use;
- self-harm and suicide-related harm, addressed in detail in Section 4.1.8 with elevated severity classification (serious harm industry-wide with case-by-case deviation up to Level 2 prioritisation);
- eating-disorder-related harm, addressed in detail in Section 4.1.8;
- amplification of distress where users encountering content engaging wellbeing concerns have inadequate access to in-app support resources.

The contemporary research base – including the US National Academies (2024), Norwegian Screen Time Committee (2024), Belgian Superior Health Council (2025), French ANSES (December 2025), and other authoritative reviews referenced above – supports a calibrated approach that recognises both the legitimate concerns documented across the research and the methodological complexity that has prevented the field from reaching settled population-level causal conclusions. The depth and durability of the harms identified in the research base is, however, sufficient to support a severity classification of severe harm industry-wide, consistent with prior Reports.

DSA Risk Factors

Consistent with Article 34(2) DSA, the following service and general risk factors have been considered in relation to this risk on Snapchat's in-scope services.

Design of recommender systems and other algorithmic systems	Recommender systems on the in-scope services do not, in the ordinary course, surface content patterns associated in academic and policy research with negative effects on physical and mental well-being. The Content Guidelines for Recommendation Eligibility set additional standards for content eligible for algorithmic distribution, with particular attention to wellbeing-relevant categories. The structural absence of features identified across the research as engagement-maximising drivers of compulsive use – autoplay, algorithmically curated feeds for unvetted user-generated content, follower-count gamification – is a foundational mitigation. Further detail in Section 5.6.

[redacted]

Content moderation systems	Substantive moderation framework recorded in Section 4.1.6 (Adult Sexual Content), Section 4.1.7 (Harassment and Bullying), Section 4.1.8 (Self-Harm and Suicide), Section 4.1.9 (Violent or Dangerous Behaviour), Section 4.1.10 (Harmful False Information), and Section 5.4. The Proactive Safety Detection System for severe-harm content (including self-harm) operates at upload. Reactive reporting is available through the in-app reporting flow with dedicated reporting reasons across the wellbeing-relevant categories. Further detail in Section 5.4.
Applicable terms and conditions and enforcement	The Community Guidelines on Self-Harm, Suicide and Eating Disorders (Section 4.1.8), on Sexual Content, on Harassment and Bullying, on Threats, Violence and Harm, and on Harmful, False or Deceptive Information together cover the substantive prohibitions most directly engaging wellbeing harms. The Community Guidelines apply uniformly across in-scope services and equally to all users. Further detail in Section 5.2 and Section 5.5.
Systems for selecting and presenting advertisements	Specific wellbeing-relevant advertising measures include: the prohibition on advertising-targeting based on minor users' activity data; the [redacted]; the prohibition on advertising for vapes/e-cigarettes (recorded in Section 4.1.3); [redacted] the rejection of advertisements containing unsubstantiated claims; the discontinuation of EU political advertising as of October 2025. Further detail in Section 5.7.
Data-related practices	PASS Review and DDIA framework apply to product and feature changes that may bear on physical and mental well-being. Digital-well-being review is incorporated as a structural component of product design (referenced across the DDIA documentation for Snaps and Chats, Family Center, Lenses, Snap Map, Public Profiles, Public Content, Private Stories, and Generative AI features). [redacted] Further detail in Section 4.2.4 and Section 6.3.

[redacted]

Intentional manipulation (Article 34(2))	[redacted]
Amplification and rapid, wide dissemination	Amplification of wellbeing-relevant harmful content is structurally constrained by the platform architecture (recommended content in public feed is moderated; no broad reshare functionality of Stories; short default retention for public stories (unless saved to Public Profile); pre-distribution moderation for content reaching wide audiences; user-generated content does not automatically advance from one user's content to the next user's content, or follower-count gamification). [redacted] in Section 6.6.
Regional and linguistic aspects	Member State frameworks bearing on the physical and mental well-being of teen users are increasingly active and evolving. Of particular note for the 2026 assessment cutoff: France's ANSES report on adolescents and social networks (December 2025); Belgium's Superior Health Council report on screen time (2025); Norway's Screen Time Committee report (2024); ongoing implementation of national Age Appropriate Design Codes; and a wider EU-level engagement on minor protection through the Article 28 DSA Guidelines. Snap engages with Member State authorities, civil-society organisations focused on youth well-being, and academic researchers on a continuing basis.

Overall Potential Risk Prioritization

Under the Section 3.5 matrix, this risk would ordinarily be Level 3. Snap has nevertheless assessed it at Level 1 on a case-by-case basis, in light of the severity analysis above and the following considerations:

[redacted]

- the depth and durability of potential wellbeing harm to affected users, particularly minors and users with pre-existing vulnerabilities;
- the recognised vulnerability of the minor-user population [redacted] of Snapchat's average monthly active recipients in the European Union are in the 13-17 cohort, representing approximately [redacted] of the EU user population) to wellbeing-relevant exposure pathways, supported by the EU Charter Article 24 elevated treatment of children's rights and by the Article 28 DSA framework;
- the contemporary regulatory and policy attention to platform-wellbeing matters, including through the Article 28 Guidelines, Member State frameworks [redacted] and the broader EU/EEA regulatory environment;
- the structural integration of wellbeing-relevant conduct across multiple Category 1 underlying categories, supporting an integrated mitigation framework rather than a single-category response.

This deviation pattern was recorded in the 2025 Report and is preserved in the 2026 Risk Assessment. There is no change in this determination from the 2025 Report. The deviation is consistent with the case-by-case treatment of other risks deviating up from the matrix default (4.1.1 CSAM, 4.2.6 Children's Rights, and 4.4.3 Minors all deviating up to Level 1; 4.1.8 Self-Harm, 4.2.1 Human Dignity, and 4.4.2 Gender-Based Violence deviating per the determinations recorded in those Sections).

[redacted]

Snap's Mitigations

The mitigations in place for this risk are mapped to Article 35(1)(a)-(k) DSA below. The framework operates principally through: (i) the structural platform-architecture features that distinguish Snapchat from peer platforms in respects most directly relevant to wellbeing (user-generated content does not automatically advance from one user's content to the next user's content, recommended content in public feed is moderated, no follower-count gamification, short default

[redacted]

retention for public stories (unless saved to Public Profile), light-touch engagement mechanics under digital-well-being review); (ii) the underlying Category 1 mitigation sets covering wellbeing-relevant content categories; (iii) Family Center oversight; (iv) Here for You in-app educational resources; (v) the Digital Well-Being Index research framework (Section 6.6); (vi) the integrated DDIA framework operationalising digital-well-being review at the product-design level; and (vii) Safety Advisory Board consultation. Descriptions are illustrative; detailed descriptions are provided in Section 5.

Snapchat Design and Function - Art. 35(1)(a)	Applies. The Snapchat platform architecture is materially differentiated from peer platforms in respects most directly relevant to physical and mental well-being: (i) recommended content in public feed is moderated; (ii) user-generated content does not automatically advance from one user's content to the next user's content; (iii) no follower-count gamification at the structural level; (iv) short default retention for public stories (unless saved to Public Profile); (v) bidirectional friending requirement; (vi) no broad reshare functionality of Stories; (vii) pre-distribution moderation for content reaching wide audiences; (viii) light-touch engagement mechanics (Streaks, Best Friends) calibrated under digital-well-being review. The minor-account default settings (Section 5.1.3) strengthen the framework for minor users specifically. Further detail in Section 5.1.
Terms and Enforcement - Art. 35(1)(b)	Applies. Substantive Community Guidelines prohibitions covering wellbeing-relevant categories recorded in Section 4.1.6, 4.1.7, 4.1.8, 4.1.9, and 4.1.10. The Severe Harms explainer addresses content engaging imminent threats to physical or emotional well-being. The Sensitive Content Policy adjusts the eligibility of content for algorithmic distribution where wellbeing-relevant considerations engage. Further detail in Section 5.2 and Section 5.5.
Moderation - Art. 35(1)(c)	Applies. Substantive moderation framework recorded in the underlying Category 1 Sections and in Section 5.4. The Proactive Safety Detection System operates against severe-harm wellbeing content (including imminent self-harm). [redacted] Further detail in Section 5.4.

Algorithmic Systems - Art. 35(1)(d)	Applies. Algorithmic systems do not categorise or recommend content engaging wellbeing harm; no corresponding interest categories exist. The Content Guidelines for Recommendation Eligibility set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution. Further detail in Section 5.6.
Advertising Systems - Art. 35(1)(e)	Applies. Specific advertising measures bearing on wellbeing include: the prohibition on advertising-targeting based on minor users' activity data; the human review of health-related advertisements; the rejection of advertisements containing unsubstantiated claims; the prohibition on advertising for vapes/e-cigarettes (Section 4.1.3); the prohibition on advertising-targeting categories likely to undermine wellbeing. Further detail in Section 5.7.
Ongoing Risk Detection and Management - Art. 35(1)(f)	Applies. The Section 6 ongoing risk-detection processes monitor wellbeing-relevant signals through Prevalence Testing across the underlying Category 1 categories, Digital Well-Being Index survey research, Safety Advisory Board consultation, engagement with civil-society organisations focused on youth well-being and digital health, and engagement with academic researchers studying platform-wellbeing relationships. Further detail is provided in Sections 6 and 6.6.
Trusted Flaggers - Art. 35(1)(g)	Applies. Notices submitted by designated trusted flaggers under Article 22 DSA may concern the underlying content categories that bear on physical and mental well-being, including self-harm, harassment, sexual content and violent or dangerous behaviour. Those notices are processed through the applicable priority-notice procedures and considered alongside the broader risk-detection inputs described in Section 6. Further detail is provided in Section 5.10.
Codes and Crisis Protocols - Art. 35(1)(h)	Applies. Snap's Content Crisis Response Protocol (Section 5.12) is available for activation in events presenting elevated wellbeing risk. Snap engages with civil-society organisations focused on youth well-being (including organisations that operate digital-wellness training and education programmes) and with the Commission and Member States on the Article 28 Guidelines (recorded in detail in Section 5.8). Further detail in Section 5.11.

Transparency - Art. 35(1)(i)	Applies. The Community Guidelines, the Severe Harms explainer, the Safety Center, the in-app educational resources of Here for You (covering anxiety, depression, eating disorders, grief, identity, bullying, and other wellbeing-relevant topics with curated content from established mental-health organisations), the parents site, and Snap's Transparency Reports together provide transparency on the framework. Snap publishes its Digital Well-Being Index research findings on a continuing basis. Further detail in Section 5.3.
Protection of Minors - Art. 35(1)(j)	Applies. The full minor-protection framework is recorded in Section 5.8 (Protection of Minors). Mitigations specific to minors that bear on wellbeing include: (i) the contact-restriction defaults (Section 5.1.3); (ii) Family Center oversight, including content-restriction settings, contact-visibility settings, and the conversation tools that enable parent-teen wellbeing dialogue; (iii) the absence of advertising-targeting based on minor users' activity data; (iv) the integration of the AADC Risk Assessment within the DDIA framework; (v) the Here for You in-app resources accessible to teen users encountering wellbeing-relevant search queries or content; (vi) the minor-account default settings and onboarding refinements. Further detail in Section 5.8.
Content Authenticity - Art. 35(1)(k)	Applies. The Community Guidelines apply without regard to whether wellbeing-relevant content is user-generated or AI-generated. Generative-AI features are subject to specific PASS Review and digital-well-being review, including in respect of input/output scanning and retry logic for unsafe outputs. Lens AR features designed to alter user appearance are subject to design review including consideration of body-image and self-esteem implications. Further detail in Section 5.9.

Conclusion

On the basis of the likelihood and severity analysis, the Article 34(2) factors, and the mitigations identified above Snap has concluded that the mitigations in place against serious negative consequences on physical and mental well-being on Snapchat's in-scope services are reasonable, proportionate and effective.

Snapchat's platform architecture is materially differentiated from peer platforms in respects most directly relevant to physical and mental well-being. The absence on the in-scope services of the design features most consistently identified across the platform-wellbeing research base as drivers of compulsive use and negative wellbeing outcomes, autoplay on user-generated

[redacted]

content, open recommender feeds for unvetted content, follower-count gamification, is a foundational structural mitigation. [redacted]

The framework operates through the integrated Category 1 mitigation sets covering the wellbeing-relevant content categories (Sections 4.1.6, 4.1.7, 4.1.8, 4.1.9, 4.1.10), supplemented by Family Center oversight, the Here for You in-app educational resources, the Digital Well-Being Index research framework (Section 6.6), the integrated DDIA framework operationalising digital-well-being review at product design, and the Safety Advisory Board consultation. The minor-protection refinements, the minor-account default settings, the onboarding redesign, the Family Center accessibility elevation, the friending-protection refinements, the profile-level moderation expansion, and the warning JIT notices for contact from accounts in elevated-scamming regions, strengthen the framework across the identified operational dimensions.

The contemporary research and regulatory base - including the US National Academies (2024), the Norwegian Screen Time Committee (2024), the Belgian Superior Health Council (2025), the French ANSES report (October 2025), the Nature Reviews Psychology review by Orben et al. (2024), and other authoritative bodies - has materially evolved during 2025 toward a more nuanced view of the platform-wellbeing relationship. The 2026 Risk Assessment preserves Snap's calibrated framing of this evolving research and regulatory environment: recognising both the legitimate concerns documented across the research base and the methodological complexity that has prevented the field from reaching settled population-level causal conclusions. Snap's mitigation framework is calibrated accordingly, with attention to specific use patterns and identified risk pathways rather than to generic time-based or access-based restrictions.

There is no change in the overall conclusion from the 2025 Report. The mitigation framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, including continued Digital Well-Being Index research, Safety Advisory Board consultation, engagement with civil-society organisations focused on youth well-being, engagement with academic researchers, engagement with Member State frameworks [redacted] and engagement with the Commission and Member States on the Article 28 Guidelines implementation, with particular attention to: the operational outputs of the minor-protection refinements; the continued evolution of the external research and regulatory base; and the wellbeing implications of generative-AI features as those capabilities develop.

5. Specific Mitigations

Article 42(4)(b) DSA requires providers of Very Large Online Platforms and Very Large Online Search Engines to transmit and make publicly available the specific mitigation measures put in place pursuant to Article 35(1) DSA. Article 35(1) DSA requires providers of Very Large Online Platforms and Very Large Online Search Engines to put in place reasonable, proportionate and effective mitigation measures, tailored to the specific systemic risks identified pursuant to Article 34 DSA, with particular consideration to the impacts of such measures on fundamental rights. Article 35(1)(a) to (k) sets out a non-exhaustive list of categories of measures that may be applied, where relevant.

Section 4 of this Report records (i) the assessment of the specific systemic risks identified in respect of Snapchat's in-scope services; (ii) a risk-level summary of the mitigation measures tailored to each risk; and (iii) Snap's conclusion as to whether those mitigation measures, taken together, are reasonable, proportionate and effective. This Section 5 provides the detailed description of the mitigation measures, organised by the Article 35(1) categories. The detailed description is the principal vehicle by which this Report complies with the obligation in Article 42(4)(b) DSA.

Article 35(1) to Section 5 mapping

The table below maps each category in Article 35(1)(a) to (k) DSA to the Section of this Report in which the corresponding mitigation is described in detail. Article 35(1)(f) - reinforcement of internal processes, resources, testing, documentation and supervision as regards detection of systemic risk - is described in Section 6 (Ongoing Risk Detection and Management), consistent with the structure of prior Reports.

Art. 35(1)	Mitigation category	Report section
(a)	Adapting the design, features or functioning of services, including online interfaces	**Section 5.1 - Snapchat Design and Function**
(b)	Adapting terms and conditions and their enforcement	**Section 5.2 - Terms; Section 5.5 - Enforcement**
(c)	Adapting content moderation processes, including speed and quality of processing notices	**Section 5.4 - Content Moderation**
(d)	Testing and adapting algorithmic systems, including recommender systems	**Section 5.6 - Algorithmic Systems**
(e)	Adapting advertising systems and presentation of advertisements	**Section 5.7 - Advertising Systems**

(f)	Reinforcing internal processes, resources, testing, documentation and supervision for detection of systemic risk	**Section 6 - Ongoing Risk Detection and Management**
(g)	Cooperation with trusted flaggers (Art. 22) and implementation of decisions of out-of-court dispute settlement bodies (Art. 21)	**Section 5.10 - Trusted Flaggers and Section 5.11 - Dispute Settlement Bodies**
(h)	Cooperation through codes of conduct (Art. 45) and crisis protocols (Art. 48)	**Section 5.12 - Codes and Crisis Protocols**
(i)	Awareness-raising and interface adaptation to give recipients more information	**Section 5.3 - Transparency**
(j)	Targeted measures to protect the rights of the child, including age verification and parental control tools	**Section 5.8 - Protection of Minors**
(k)	Distinguishing generated or manipulated content through prominent markings and functionality for recipients to indicate such content	**Section 5.9 - Content Authenticity**

Horizontal and risk-specific application

A feature of Snap's mitigation framework is that a number of measures operate horizontally across multiple systemic risks. For example, the content-moderation systems described in Section 5.4 address hash-based detection, machine-learning classifiers, human review and specialised workflows that apply to several Category 1 harms, and by extension to the derived-rights analyses in Category 2 and the derived public-health analysis in Category 4. The Article 34(2) factor analysis in Section 4 reflects this by cross-reference, rather than by exhaustive repetition.

Where a mitigation is tailored to a specific risk, for example, the hash-based detection applied to CSAM in Section 5.4, or the Heads Up interstitial applied to drug-related search terms in Section 5.3, the tailoring is recorded in the relevant sub-section, and the corresponding risk in Section 4 references the tailored measure.

Language availability

User-facing measures and supporting information are made available in the languages required by the applicable DSA provisions and, where relevant, in the languages in which Snapchat is available. In particular, Snap publishes its terms and conditions in the official languages of the Member States in which it offers Snapchat, in accordance with Article 14(6) DSA.

Each mitigation category is addressed in turn in the sub-sections that follow.

5.1 Snapchat Design and Function

5.1.1 Introduction

Design decisions have been integral to Snap's approach to mitigating systemic risk since the service's inception. This Section 5.1 describes the design and functional mitigations relied upon to support the reasonableness, proportionality and effectiveness analysis under Article 35(1) DSA, including foundational architecture decisions carried forward from prior Reports and refinements introduced during the 2026 assessment period.

The design framework comprises: (i) privacy and safety by design principles applied to new product and feature development; (ii) a Dark Patterns Design Policy prohibiting the use of deceptive design patterns and misleading nudge techniques; (iii) an Online Interface Design Process that governs development, testing, adaptation and deployment of interface changes; (iv) defined roles and responsibilities for review and approval of changes; and (v) annual review of the Dark Patterns Design Policy by the DSA Cross-Functional Governance Team. The design framework operates in combination with the mitigations described in Sections 5.2 to 5.11 and the ongoing risk detection measures described in Section 6.

5.1.2 Oversight and Administration

Roles and Responsibilities

Role	Responsibilities
Legal Teams	Define roles and responsibilities for review of changes to the design, features or functioning of Snapchat services. Review proposed feature changes prior to launch for potential dark patterns on online interfaces.
Design, Product and Engineering Teams	Responsible for adherence to Snap's Dark Patterns Design Policy when developing, testing, adapting and deploying changes to Snapchat interface and functionality.
DSA Cross-Functional Governance Team	Reviews the Dark Patterns Design Policy on an annual basis to ensure incorporation of relevant regulatory requirements.
DSA Compliance Officer	Oversees compliance of Snapchat's online interfaces and organisation with DSA requirements, including Article 25 (online interface design and organisation) and Article 35(1)(a) (design-based mitigation).

[redacted]

The Dark Patterns Design Policy is reviewed by the DSA Cross-Functional Governance Team on an annual basis, and the Privacy and Safety by Design process is reviewed on the same cadence (see Section 6.3).

5.1.3 Adaptations and Mitigations

As a result of the privacy and safety by design approach described in Section 6.3, Snapchat's in-scope services are configured with foundational design features that mitigate the risks identified in Section 4. This Section 5.1.3 records those foundational features, together with surface-specific mitigations for Spotlight, Discover, Public Profile, Snap Map, Lenses and Advertising, and the consolidated minor-account design measures. While not all features described below are in scope of the risk assessment, they are included where they bear on the overall design posture of in-scope services. This list is illustrative, not exhaustive.

Foundational design elements

The following foundational design elements have been in place since prior Reports:

- **Contact Restrictions.** As a general matter, users can direct-message one another only if they are bidirectional friends or if they are in each other's device contact book. This adds friction to contact from non-connected users and is an important mitigation against unwanted contact and solicitation.
- **Private friend lists.** Friend lists are not disclosed to other users, and the total number of friends is not exposed. This reduces the potential for non-connected users, including strangers, to navigate from one user's friend graph to another.
- **Camera-first interface.** The application opens to the Camera by default, not to a feed. This design orientation emphasises creation and direct communication over passive consumption of an algorithmic feed.
- **Control over Visibility**. Users choose how they want to share content on Snapchat. Whether content is viewable only by bidirectional friends, a custom audience, or more widely it is up to the user to decide when sharing content.
- **Absence of public vanity metrics.** Public "like" counts and similar vanity metrics are not displayed on user Stories, reducing social-comparison and validation-seeking dynamics.
- **Ephemerality by default.** Short default retention for public stories (unless saved to Public Profile).

Independent of any individual feature choice, time-spent patterns on the service continue to concentrate in private and peer-to-peer surfaces. Between 1 July 2025 and 30 June 2026, approximately [redacted] of EU time on platform was spent on the Camera and Friends tabs, comprising [redacted] on Camera and [redacted] on Friends. By comparison, approximately [redacted] was spent on Snap Map, [redacted] on Spotlight and [redacted] on the Stories tab. These figures support the conclusion that use of

[redacted]

Snapchat remains concentrated in creation and direct communication rather than public-content discovery surfaces.[120]

Account creation, age assurance and minor-account design

This sub-section consolidates the design-layer mitigations specific to minor-account experiences. [redacted] Section 4.4.3 (Negative Effects on Minors). Related child-safety and well-being consequences are addressed in Sections 4.1.1, 4.2.6 and 4.4.4, while this Section describes the corresponding design-layer mitigations.

Default settings for minor accounts

Minor accounts are configured with the following default settings and controls:

[redacted]

[120] Internal Analytics data, EU time on platform by tab, 1 July 2025 to 30 June 2026. Percentages are rounded to the nearest whole number.

[redacted]

[redacted]	[redacted]

The rationale for these defaults and controls is that the minor-account experience, as a general matter, should be configured to minimise unwanted contact, reduce discoverability by non-connected users, and preserve the minor's ability to engage with the service on terms the minor actively chooses rather than on terms the minor would be required to de-select. Further detail on Snap's assessment against the Article 28 Guidelines is provided in Section 5.8.2.

Minor-account onboarding flow

The minor-account onboarding experience remains subject to Snap's Privacy and Safety by Design process. [redacted]

Spotlight

Spotlight is a content discovery surface for user-generated Snaps. The following design measures apply:

- creators may choose to disable comments or to approve comments manually; [redacted]
- comments on content posted by minors are restricted to other minors; adults cannot comment on minors' Spotlight posts;

[redacted]

Discover

Discover presents Creator Stories, including Media Partner content and user-generated content from Snap Stars and Popular Users meeting a follower-count threshold. The following design measures apply:

- viewer comments are generally not available on Discover; where enabled, comments are auto-moderated;
- view counts are not displayed on Discover Stories;
- publication duration for Discover content is limited, subject to extensions available to Snapchat+ subscribers;
- Discover content does not automatically advance from one user's content to the next user's content; and
- users' likes and interests are not publicly displayed.

Public Profile

Public Profiles permit users to showcase public content. The following design measures apply:

- users can clear information and content added to their Public Profile in a single action; Spotlight and Snap Map posts are managed separately;
- users can individually show or hide their Spotlight content on their Public Profile, both at submission and afterwards;
- users can disable remixes of their public content;
- [redacted]
- Public Profile users can disable, filter or block Story replies, and can manage who may contact them through privacy settings;

[redacted]

- Public Profile owners do not receive notifications for new followers, reducing social-validation dynamics; and
- in-app educational content and Just-in-Time notices inform users of the public nature of content posted via Public Profiles.

Snap Map

Snap Map is an interactive location surface. The following design measures apply:

- location sharing is off by default for all users. Users must affirmatively enable location sharing;
- location can be shared only with bidirectional friends; there is no option to share location with non-friends;
- users can, at any time, revoke location sharing from the Snap Map settings;
- Place Stories and Community Stories are subject to proactive and reactive moderation (see Section 5.4); and
- view counts are not displayed on Snap Map Stories.

Lenses

Lenses are AR experiences created by Snap's internal Lens Team and by community developers. The following design measures apply:

- Lenses use object detection rather than facial identification. Lenses can distinguish whether a face is present in a frame but do not identify specific faces;
- voice-data collection for EU Snapchatters is off by default, and is used only to provide the specific service requested by the user; and
- Lenses are designed with race, gender, ethnicity and cultural considerations in mind; user feedback is taken into account, and Lenses generating a high rate of user reports are re-reviewed.

Advertising

Advertising on in-scope services is subject to design-level mitigations described in greater detail in Section 5.7. The following design measures are noted here for completeness:

- [redacted]
- as of October 2025, Snap has discontinued political advertising in the European Union, consistent with the EU's Political Advertising Regulation; see Section 5.7; and
- users in the European Union can disable personalised public content recommendations via the "European Union Controls" settings.

[redacted]

Heads Up and safety-support surfaces

Snap provides a set of in-app safety-support surfaces that complement the content moderation and enforcement mitigations described in Sections 5.4 and 5.5. The principal surfaces include:

- **Heads Up.** Heads Up is an informational safety-support surface displayed at the top of Search results when users search for predefined drug-related terms. During the assessment period, localised Heads Up portals were available for the United Kingdom, France and the Netherlands, with the Netherlands portal launched in June 2026; users elsewhere in the EU were directed to the English-language US portal. Snap continues to explore additional localised versions. Further detail on its operation and corresponding Search-related measures is provided in Sections 4.1.3 and 5.4.4.
- **Here For You.** An in-app resource presented to users engaging with search terms relating to suicide, self-harm or mental health, with curated content from clinical experts and support organisations.
- **Family Center.** A parental oversight surface that permits parents to view recent chat partners and friend additions for a consenting minor user. [redacted].

[redacted]

5.1.4 Integrations with other mitigations

Terms

Design measures are underpinned by the Terms of Service, the Community Guidelines, the Content Guidelines for Recommendation Eligibility, and product-specific terms. These are described in Section 5.2. [redacted] reflect operational implementations of underlying policy positions recorded in the Community Guidelines.

Content Moderation

The design configurations that channel content into Public, Limited Broadcast and Private content surfaces inform the content moderation approach described in Section 5.4, including the intensity

[redacted]

and type of moderation applied at each content surface and the virality circuit-breaker applied to content reaching wide distribution.

Content Distribution

The algorithmic recommendation systems that determine content distribution to wide audiences, addressed in Section 5.6, interact with design features such as the Content Guidelines for Recommendation Eligibility, the pre-distribution moderation on Spotlight and Discover, and the absence of algorithmically amplified political content on in-scope services.

5.1.5 Online Interface Design Process

[redacted]

5.1.6 Conclusion

The design and functional mitigations described in this Section operate across in-scope services and, in combination with the mitigations described in Sections 5.2 to 5.12 and the ongoing risk detection measures described in Section 6, support the Article 35(1) proportionality analysis recorded in the corresponding risk assessments in Section 4.

5.2 Terms

5.2.1 Introduction

[redacted]

The terms and policies governing the use of Snapchat establish the contractual framework within which the in-scope services operate and set the substantive behavioural expectations that underpin the moderation and enforcement mitigations described in Sections 5.4 and 5.5. This Section 5.2 sets out how Snap's terms and conditions (as defined below) and other terms and policies contribute towards Snap's mitigation measures for the purposes of Article 35(1)(b) DSA, as well as describe Snap's various feature-specific recommender systems. The corresponding enforcement mitigation layer is recorded in Section 5.5.

For the purposes of this section, references to Snap's "terms and conditions" means the following key documents:

- Terms of Service, which establish the contract between Snap and each user of the services;
- Community Guidelines, which elaborate the behavioural expectations applicable to all content and conduct on the services and are incorporated into the Terms of Service by reference
- "Snap and Ads" privacy explainer, which explains how Snap determines which ads are displayed and when.
- "How We Rank Content on Spotlight" support page, which explains how content on Spotlight is ranked and recommended to users.
- "How We Rank Content on Discover" support page, which explains how content on Discover is ranked and recommended to users.
- "How We Rank Content on Lenses" support page, which explains how Lenses are ranked and recommended to users.
- "How We Rank Content on Maps" support page, which explains how content in Maps is ranked and recommended to users.

The following terms and policies, while not forming part of Snap's "terms and conditions", are also covered in this section as they contribute to Snap's mitigation layer for the purposes of Article 35(1)(b) DSA:

- the Privacy Policy, which describes personal data processing and user rights;
- the Content Guidelines for Recommendation Eligibility, which establish the additional standards applicable to content eligible for algorithmic recommendation beyond a creator's friends or followers (see also Section 5.6);
- the Advertising Policies and the Commercial Content Policy, which apply to paid and organic commercial content respectively (see also Section 5.7);
- the Creator Monetization Policy, which applies to users participating in Snap's Creator Program; and
- product-specific terms governing the use of particular features, including Spotlight, Discover, Lenses, Advertising, and generative-AI features.

[redacted]

5.2.2 Oversight and Administration

Change Management

The terms and conditions are reviewed by Snap's Management body (including Legal and Public Policy stakeholders) on an annual basis and, where a pressing need arises, on an ad hoc basis. Reviews consider whether updates are required in light of new regulatory requirements, changes to the services, or considerations arising from the risk assessment. Where updates are identified, proposed revisions are drafted and approved by the owner of the respective document. A changelog is maintained for each of the terms and conditions to document the update and approval process.

Once finalised and approved, revised terms and conditions are localised in all supported languages, including each of the official languages of the European Union as required by Article 14(6) DSA. Where a revision is considered material, users are given reasonable advance notice of the change via an in-app notification. To the extent any changes to the terms and conditions are determined to be material or significant, Snap will provide users with reasonable advance notice.

Snap provides an in-app pop-up to notify recipients of the service of these types of changes to the terms and conditions. Refer to the 2025 Risk Assessment for an example.

Roles and Responsibilities

[redacted]

5.2.3 Terms and Conditions

Terms of Service

The Terms of Service[121] are the primary contractual document governing use of Snapchat. Key provisions include:

- **Minimum age.** The services are not directed to children under the age of 13 or, where higher, the minimum age specified in a user's particular country or region, and a user must confirm that they are that age or older to create an account and use the services. Where applicable law imposes a higher minimum age, the applicable local age applies.
- **Respecting others' rights.** Users may not use the services, or enable others to use the services, in a manner that violates or infringes the rights of publicity, privacy, copyright, trademark, or other intellectual property right of another person.

[121] https://www.snap.com/terms

- **Safety.** Users agree to comply with the Terms of Service, including the Community Guidelines and any other policies made available to maintain the safety of the services.
- **Content moderation.** Much of the content on the services is produced by users, publishers and third parties; the content is the responsibility of the user or entity that submits it. Snap reserves the right to review, moderate or remove content that appears on the services, though it does not review all content. Users may report content and accounts for violation of the Terms of Service or other policies.
- **Service restriction, termination and suspension.** Snap may restrict, terminate or temporarily suspend access to the services where a user fails to comply with Snap's terms or policies, Community Guidelines, or applicable law, for reasons outside of Snap's control, or for any other reason. Where reasonably possible, Snap seeks to give reasonable advance notice. In deciding to restrict, terminate or suspend, Snap takes into account all relevant facts and circumstances apparent from the information available.
- **Notice and action; appeals.** The Terms of Service reference the mechanisms through which users may submit notices regarding content and accounts, and through which affected users may seek review of enforcement decisions. The detailed operation of those mechanisms, including compliance with Articles 16 (notice and action), 17 (statement of reasons), 20 (internal complaint-handling) and 21 (out-of-court dispute settlement) DSA, is described in Section 5.5.

Community Guidelines

The Community Guidelines are incorporated into the Terms of Service by reference and elaborate the behavioural expectations applicable to all content and conduct on the services. The Guidelines are organised into six categories, each covered by a dedicated explainer in the Community Guidelines Explainer Series (available in the Annex):

Category	Scope (summary)
Sexual Content	Prohibits: sexual exploitation or abuse of a minor (including CSAM, grooming for sexual purposes, sexualisation of children); communication or behaviour intended to persuade, trick or coerce a minor for sexual abuse or exploitation; all forms of sexual exploitation, including sex trafficking and sextortion; non-consensual intimate imagery (NCII); sexual harassment; pornographic content; and offers of sexual services.
Harassment and Bullying	Prohibits: targeted harassment, bullying, stalking, doxxing, and threats directed at individuals or groups, including across accounts.

Threats, Violence & Harm	Prohibits: encouraging or engaging in violent or dangerous behaviour; intimidation or threats directed at persons, groups, or property; gratuitous or graphic violence, including animal abuse; glorification of self-harm, including promotion of self-injury, suicide, or eating disorders.
Harmful False or Deceptive Information	Prohibits: spreading false information that causes harm (including denial of tragic events, unsubstantiated medical claims, and content undermining the integrity of civic processes); impersonation; spam and follower-growth schemes; fraud and deceptive practices.
Illegal or Regulated Activities	Prohibits: any use of Snapchat for content that is illegal in the user's jurisdiction or for any illegal activity, including promoting, facilitating or participating in the sale of illegal or regulated drugs, contraband (including CSAM), weapons, counterfeit goods or documents; facilitation of exploitation, including sex trafficking, labour trafficking and other human trafficking; unauthorised promotion of regulated goods or industries, including gambling, tobacco and vape products, and alcohol.
Hateful Content, Terrorism and Violent Extremism	Prohibits: use of the platform by terrorist organisations, violent extremists and hate groups; content advocating or advancing terrorism or violent extremism; Hate Speech and content that demeans, defames, or promotes discrimination or violence on protected grounds (including race, colour, caste, ethnicity, national origin, religion, sexual orientation, gender, gender identity, disability, veteran status, immigration status, socio-economic status, age, weight, or pregnancy status); use of hate symbols.

The Community Guidelines apply irrespective of whether content is user-generated or AI-generated. No material changes to the Community Guidelines categories described above were introduced during the 2026 assessment period. A subsequent update concerning the eligibility of certain AI-generated content on Spotlight took effect after the 30 June 2026 assessment cutoff and is therefore outside the scope of this Report.[122]

Recommender System Pages

[122] https://newsroom.snap.com/rewarding-authentic-creativity-on-spotlight.

The Support Site provides product- and feature-level support pages that form part of the terms and conditions, including the following ranking-methodology documents referenced in the Terms of Service in connection with recommender systems (see also Section 5.6):

- How We Rank Content on Spotlight[123],
- How We Rank Content on Discover[124],
- How We Rank Content on Lenses[125]; and
- How We Rank Content on Maps[126].

Other Support Site articles referenced from the Terms of Service or from feature-specific materials include privacy FAQs, privacy-settings guides, the Generative AI on Snapchat support page, the My Selfie Help Center article, and the "Your Privacy When You Post Publicly on Snapchat" article.

Recent updates

No significant changes were made to the terms and conditions during the 2026 assessment period.

Other relevant terms and policies

Privacy Policy

Snap also publishes a Privacy Policy which describes personal data processing in connection with the services, including the categories of personal data collected, the purposes of processing, the legal bases relied upon, the categories of recipients, the data retention approach, and the rights available to data subjects. The Privacy Policy references the Privacy Center and the Privacy by Product pages, which provide more accessible summaries of the underlying practices (described further in Section 5.3.1).

Product-Specific Terms

In addition to the terms and conditions described in detail above, Snap also has specific, publicly-available terms and policies that govern the use of additional aspects of Snapchat's features which we have referenced below for background purposes:

Surface	Applicable terms

[123] https://help.snapchat.com/hc/en-us/articles/8961653169940-How-We-Rank-Content-on-Spotlight.
[124] https://help.snapchat.com/hc/en-us/articles/8961631424020-How-We-Rank-Content-on-Discover.
[125] https://help.snapchat.com/hc/en-us/articles/34360182512916-How-We-Rank-Lenses-on-Snapchat.
[126] https://help.snapchat.com/hc/en-us/articles/35157584066964-How-We-Rank-Content-on-Maps.

Spotlight	Creators contributing content to Spotlight agree to the Creator Monetization Policy. The Content Guidelines for Recommendation Eligibility establish the policy, technical and legal requirements applicable to content submitted for algorithmic recommendation, and incorporate the Community Guidelines by reference. The Snap Spotlight Submission and Revenue Terms address the public nature of submitted content and applicable third-party rights.
Discover	Publisher and creator partners on Discover operate under individualised publishing agreements requiring compliance with the Terms of Service and Community Guidelines.
Lenses	Lens Studio Terms govern the development and submission of Lenses by independent developers. Lens Studio Submission Guidelines apply to Lens submissions. The Generative Lens Tools Terms govern the use of generative-AI tools within Lens Studio, including use of prompts and outputs and applicable rights and restrictions.
Advertising	Advertisers must accept the Snap Advertising Policies, including the obligation for advertisements to comply with applicable law and rules in the European Union and each Member State in which the advertisement is delivered. Additional requirements and safeguards apply to categories identified as higher risk, including political, health-related, complex financial products, and gambling or gaming advertising.

5.2.4 Accessing Terms and Conditions

All terms and conditions are accessible via internet search, are machine-readable, and where necessary are supported by shorter user-facing summaries (for example, the Terms of Service).

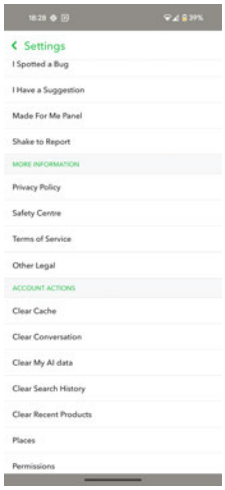
The Terms of Service and the Community Guidelines are are prominently accessible to users in the following locations:

- Persistent footer on Snap.com website:



- Link to Terms of Service in mobile app “Settings”:

[redacted]



The four recommender system pages that form part of the terms and conditions are linked to prominently within the Terms of Service.

Material changes to the Terms of Service are communicated through in-app notifications, as described in Section 5.2.2.

5.2.5 Languages

All terms and conditions are available in each of the official languages of the European Union, as required by Article 14(6) DSA and Recital 48. Other principal policy documents, product-specific terms and support materials are localised across the languages in which they are made available, having regard to the intended audience and applicable legal requirements.

5.2.6 Readability

The user base of Snapchat in the European Union is composed across a range of age groups. During the 2026 assessment period, the 18-24 cohort comprised [redacted] average monthly active recipients, representing approximately[redacted] of the EU user population, while the 13-17 cohort comprised [redacted] recipients, representing approximately 1[redacted]

Terms and conditions are designed to meet the DSA requirement that key terms be presented in a concise, clear and easily understandable form, in a publicly available and machine-readable format. The user-facing summaries are supplemented by the iconography and age-appropriate presentation used across the Privacy, Safety and Policy Hub, informed by guidance from safety experts and by the Age Appropriate Design Code. The approach is intended to make the substantive content of the Terms of Service accessible to the full range of the user base, including minor users.

[redacted]

5.2.7 Conclusion

The terms and conditions and other relevant policies and terms described in this Section establish the contractual framework and behavioural expectations that underpin the moderation and enforcement mitigations described in Sections 5.4 and 5.5. The terms and conditions are published in the official languages of the European Union as required by the DSA, are machine-readable and are supplemented by user-facing summaries and explainers. The continued operation of this framework during the 2026 assessment period, including the generative-AI-specific provisions effective from 7 April 2025 and the supporting Community Guidelines, policies and user-facing explainers, supports the conclusion that the terms-layer mitigations are reasonable, proportionate and effective for the risks to which they apply.

5.3 Transparency

This Section 5.3 describes the transparency measures considered as part of the Article 35(1) mitigation framework. The measures are designed to give users information relevant to the service, applicable rules, safety and privacy considerations, and available controls at a level of detail appropriate to the context. Information is delivered through three complementary categories:

- information provided on the Snap website and support pages (Section 5.3.1);
- information provided in app stores (Section 5.3.2); and
- information provided in the application itself, including onboarding notices, just-in-time notices, and thematic awareness surfaces (Section 5.3.3).

The transparency measures are provided in supported languages according to the applicable legal requirements, intended audience and deployment scope, as described in Section 5.3.4. [redacted] Transparency therefore operates as a supplementary mitigation alongside the design, moderation, enforcement, algorithmic, advertising, minor-protection and content-authenticity mitigations described elsewhere in Section 5.

5.3.1 Information provided on the website

Information provided outside the application is anchored in two primary surfaces: the Privacy, Safety and Policy Hub and the Support Site.

Privacy, Safety and Policy Hub

The Privacy, Safety and Policy Hub was launched in 2022, combining the Privacy Center, Safety Center and Transparency Center under a single umbrella, with a dedicated Policy Hub added in

[redacted]

2024. The purpose of the combined Hub is to reflect the substantive overlap between these areas and to provide a single navigational entry point for different audiences - users, parents, teachers, journalists, trusted flaggers, law enforcement, regulators and NGOs. The top navigation provides access to Policy, Privacy, Safety and Transparency resources, as well as the News page.

Policy Center

The Policy Center provides centralised access to Snap's substantive policies, including the Community Guidelines, the Content Guidelines for Recommendation Eligibility, the Advertising Policies, the Commercial Content Policy and the Creator Monetization Policy. The policies are presented in a shorter, topic-organised format, replacing earlier text-heavy presentation.

Privacy Center

The Privacy Center supplements the Privacy Policy with more digestible, topic-based materials: Privacy Principles (Snap's general approach to privacy), "How we use your information" (descriptions of specific data-use categories), Snap and Ads (advertising practices) and Privacy by Product (product-specific information). The Privacy by Product page provides feature-level overviews - for example, Snaps & Chats, Spotlight, Lenses, Stories and My AI - each with hyperlinks to more detailed support materials. The Privacy Center includes a video overview of Snap's privacy practices and uses visual and layered presentation intended to make the materials easier to navigate and understand.

Safety Center

The Safety Center aggregates safety resources, including: guidance on reporting content and accounts; information on the Trusted Flagger Program, the Safety Advisory Board, and the Digital Well-Being Index; Safety Policies; information for parents and information for law enforcement; and a dedicated page on Financial Sextortion. Within the Safety Center, dedicated pages are provided for Reporting (summarising reporting pathways and linking to the Reporting Quick Guide and web reporting form) and for Safety Resources and Support (including links to MindUp, the Here For You tool, country-specific helplines, and pages on sexual risks and harms).

The Safety Resources and Support pages link users to wellbeing resources, support organisations and, where available, country-specific helplines. These materials supplement the in-app Here For You resources and other safety-support surfaces described in Section 5.3.3.

Family Safety Hub and Parents microsite

The Parents microsite provides information targeted at parents and caregivers, including Snapchat 101 materials, an overview of safeguards for teens, and dedicated information on Family Center and Snap Map safety. The microsite is designed for caregivers who may not themselves use Snapchat or be familiar with the service's features.

[redacted]

Transparency Center

The Transparency Center provides access to Snap's Transparency Reports and EU-specific information published in accordance with the DSA. The EU transparency page includes information concerning average monthly active recipients in the European Union, Snap's legal representative in the European Union, procedures for EU law-enforcement requests, and the regulatory authorities overseeing Snap under the DSA.

News Page

The Hub's News page is used to publish updates on developments relevant to privacy, safety and transparency on Snapchat, including product changes, policy updates and industry engagement. The News page is an additional channel complementing Transparency Reports and product-level documentation.

Generative AI transparency

Support pages relating to generative AI features on Snapchat are published as part of the Support Site structure. The "Generative AI on Snapchat" support page provides an overview of the technology used, describes how user input data is processed, and identifies applicable transparency and safety measures, including contextual identification or watermarking for specified AI-generated outputs, reporting tools and moderation filters. The page also explains the restrictions applicable to removing or obscuring Snap watermarks from AI-generated images Snap's Terms of Service.

Feature-specific support materials include the My Selfie Help Center article, which explains how a user can set up, control and delete My Selfie data. Further detail on content-authenticity measures relating to generative AI is recorded in Section 5.9.

5.3.2 Information provided in app stores

Before a user installs the Snapchat application, supplementary information is made available through the relevant listings on Apple's App Store and Google Play. The listings include: a description of the service; data safety disclosures (as required by the respective app store policies); the minimum age requirement (currently 13 in most jurisdictions, subject to applicable law); content and privacy labels required by each platform; and the Snap Privacy Policy and Terms of Service by link. The app stores also surface user reviews and ratings, which Snap monitors alongside other user feedback channels. Snap reviews the accuracy of its app-store disclosures through applicable product and privacy change-management processes. App-store disclosures supplement, but do not replace, Snap's own legal notices and in-product transparency.

5.3.3 Information provided in the application

[redacted]

In-application information is provided through three principal mechanisms: the onboarding process, Just-in-Time (JIT) notices, and thematic awareness surfaces.

Onboarding process

The onboarding process is the principal moment at which a new user is introduced to core service features, safety settings and privacy controls. The onboarding flow includes, among other elements: date of birth entry (used as the basis for age-based feature gating, as described in Section 5.8); agreement to the Terms of Service; prominent presentation of and access to the Privacy Policy; separate permission or consent requests where applicable; introductory surfaces for contact permissions and discoverability; and introductory information on key privacy defaults (including the Ghost Mode default for Snap Map).

[redacted]

Just-in-Time (JIT) notices

JIT notices are in-application notifications delivered when a user is about to take, or has received, an action that may bear on a safety or privacy consideration. JIT notices are designed to provide concise, contextual information at the moment of use, rather than requiring the user to locate separate policy documentation.

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]

[127] Snap Inc., '[Launch doc] [EU] Default selected date to today with 24h cooldown on Birthday page During Reg' (22 June 2026); Snap Inc., 'Study Launch Document – REGISTRATION_AGE_GATING_V2' (study conducted 28 May–7 June 2026; EU rollout on iOS and Android completed 22 June 2026).

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

[redacted] Information concerning the relevant categories of data and their use is provided through Snap's layered privacy disclosures. Operational details are limited where disclosure could facilitate circumvention or undermine user safety.

[redacted]

Thematic awareness and notices

Thematic awareness surfaces deliver topic-specific information in response to user behaviour or at moments of elevated risk. The principal surfaces are:

[redacted]

- **Here For You.** An in-app resource presented when a user searches for terms related to mental health, anxiety, depression, stress, suicidal thoughts, grief or bullying. [redacted]
- **Heads Up.** An informational safety-support surface displayed at the top of Search results when users search for predefined drug-related terms. [redacted]
- **In-app safety prompts and contextual symbols.** Contextual symbols are used to identify Snap-generated AI content in certain surfaces, supporting transparency regarding content authenticity (see Section 5.9). Sensitive-topic prompts are displayed in connection with certain search behaviours or content interactions.
- **Family Center entry points.** In-application entry points allow users to locate Family Center through Settings and other relevant surfaces. Further detail on Family Center and its associated minor-protection measures is provided in Sections 5.1.3 and 5.8.

5.3.4 Languages

The Terms of Service are available in each of the official languages of the European Union, as required by Article 14(6) DSA and Recital 48. The Privacy Policy, Community Guidelines and other primary policy documents are available in a corresponding range of languages. Support pages, including product-specific support content (for example, the “Generative AI on Snapchat” support page and the My Selfie Help Center article), are localised across the languages in which Snapchat is available.

In-application transparency surfaces, including onboarding notices, JIT notices and thematic awareness surfaces, are presented in supported languages according to the relevant feature and deployment scope. Further detail on the Heads Up language expansion is provided in Section 5.1.3.

5.3.5 Conclusion

The transparency measures described in this Section operate in combination with the design, moderation, enforcement, algorithmic, advertising, minor-protection and content-authenticity mitigations described elsewhere in Section 5. The measures are designed to present the right information, at the right level of detail, at the moment of use - and to be accessible to users,

[redacted]

parents, researchers, regulators and other stakeholders through the Privacy, Safety and Policy Hub and related surfaces.

The website, app-store and in-application transparency measures described in this Section support the conclusion that the transparency-layer mitigations are reasonable, proportionate and effective for the risks to which they apply.

5.4 Content Moderation

5.4.1 Approach

This Section 5.4 describes the content moderation mitigations relied upon to support the Article 35(1) proportionality analysis recorded in Section 4. The content moderation layer operates in combination with the design mitigations in Section 5.1, the Terms framework in Section 5.2, the transparency measures in Section 5.3, the enforcement mitigations in Section 5.5, the algorithmic system adaptations in Section 5.6, and the advertising system moderation in Section 5.7.

[redacted]

- [redacted]
- [redacted]
- [redacted]

All content is subject to the Community Guidelines and the Terms of Service. Content eligible for algorithmic recommendation beyond a creator's friends or followers is additionally subject to the Content Guidelines for Recommendation Eligibility. Content is assessed proactively (through automated detection) and reactively (upon user report), using a combination of automated review and human review. [redacted]

5.4.2 Oversight and Administration

Roles and Responsibilities

Role	Responsibilities

[redacted]



[redacted]

Moderator wellness

Moderators reviewing sensitive or distressing content are treated as digital first responders and supported accordingly. [redacted]

Automated Moderation

[redacted]

Snap operates an automated moderation stack combining [redacted] The full detection-system suite is set out in Section 5.4.3.

Training, Testing and Quality Assurance

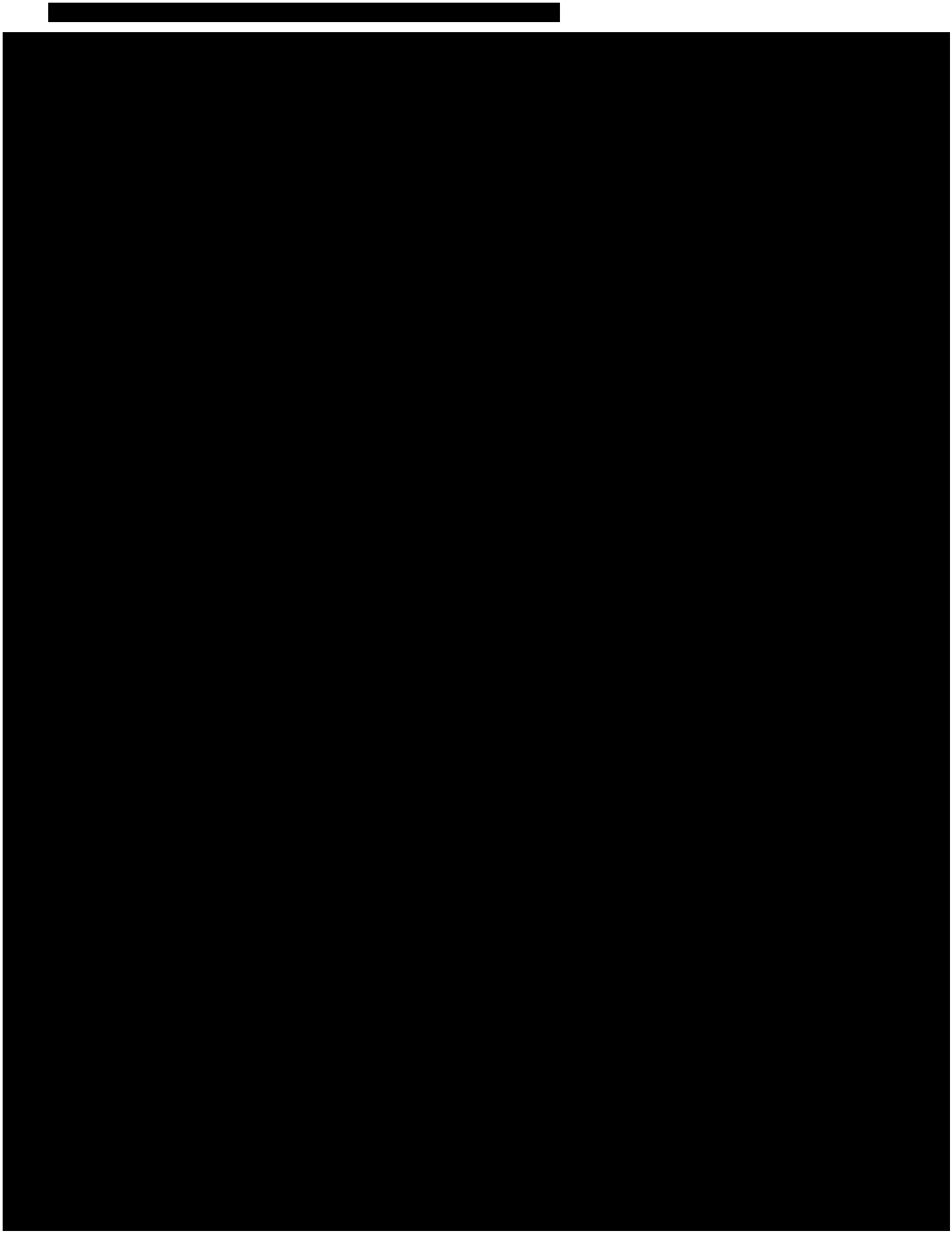
Automated moderation technology is tested and approved prior to deployment and prior to material updates. [redacted]

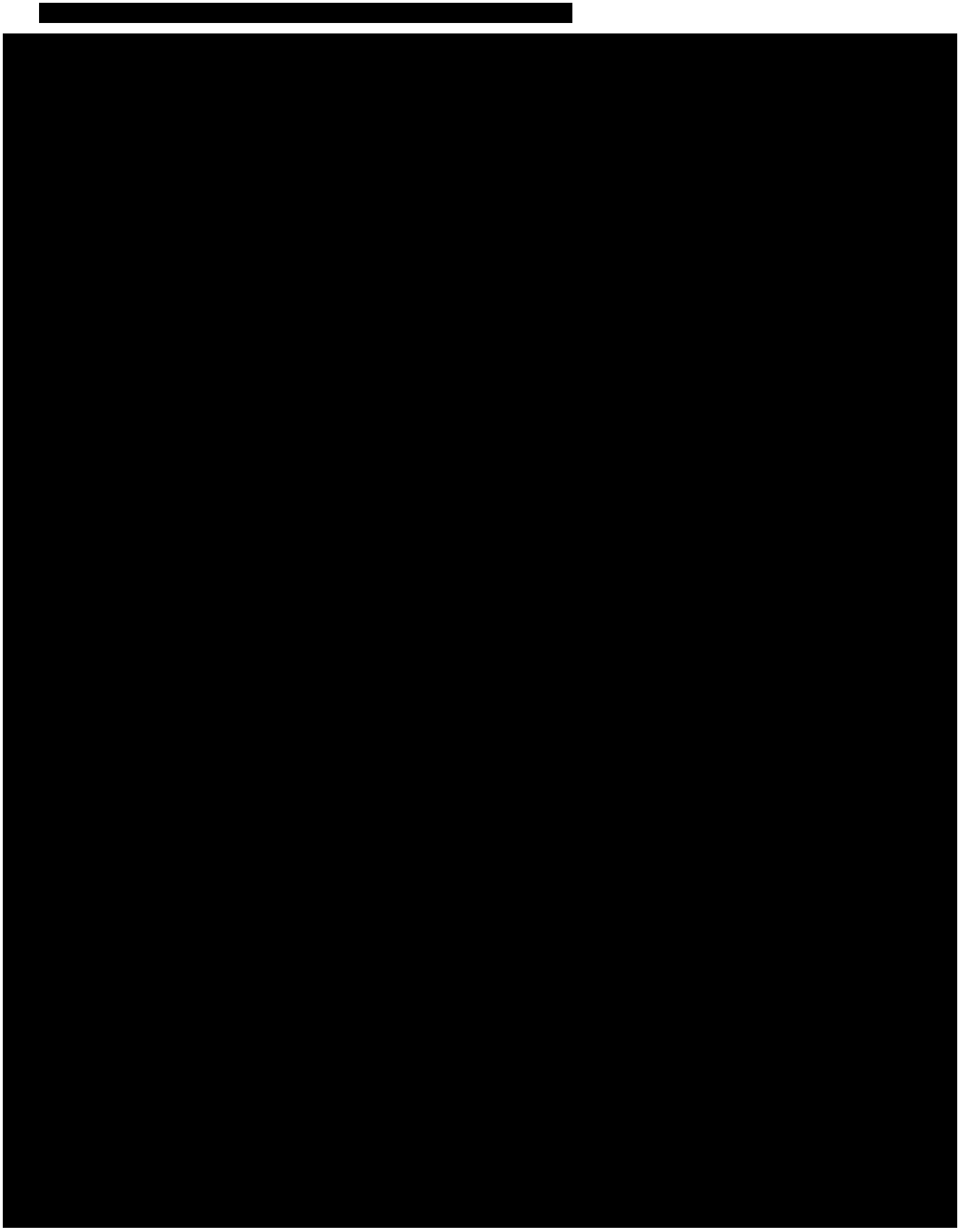
[redacted]

5.4.3 Proactive Moderation

[redacted]







[redacted]

[redacted]

[redacted]

Abusive Language Detection

ALD is a text-moderation system scanning for abusive text in [redacted]

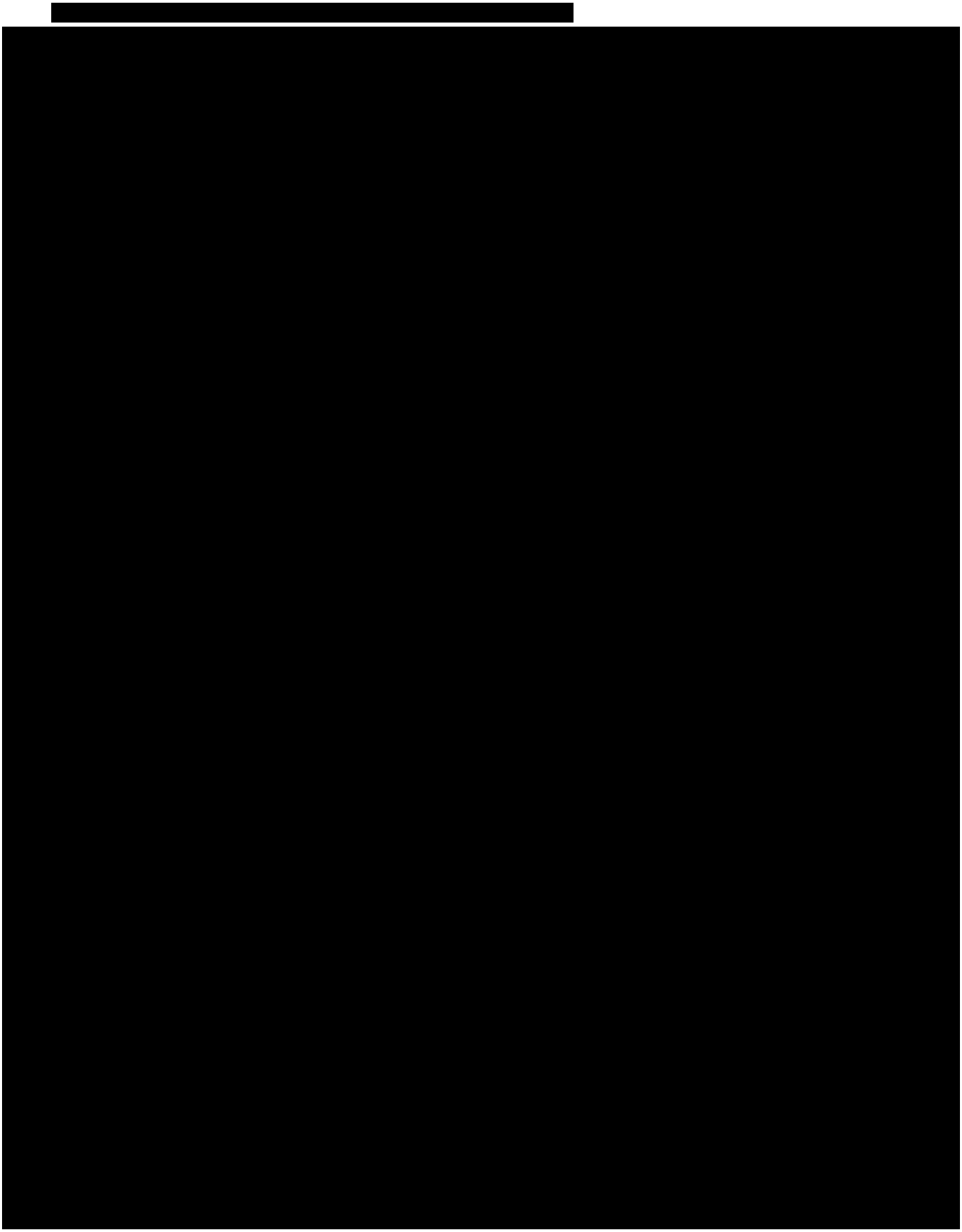
[redacted]

[redacted]

[redacted]

[redacted]

Surface / mechanism	Mitigation and effectiveness



[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

5.4.6 Reactive Moderation (Reporting)

Reactive moderation operates through in-app reporting, web-based reporting and out-of-app pathways. [redacted]

Content-level in-app reporting

Users can report individual Snaps, Stories, Spotlight posts, Discover content, Snap Map content, Lenses, Messages and other content surfaces through the in-app reporting menu. The menu is structured by reporting category (aligned to the Community Guidelines). Reports are routed to moderation queues organised by reason.

Account-level in-app reporting

Users can report accounts rather than individual content items through the account-reporting pathway. [redacted]

Web-based reporting

Web-based reporting is available through the Snapchat Support site, and operates for both logged-in and logged-out users. [redacted]

Illegal Content Notice (Article 16 DSA)

[redacted]

Illegal content is violative of the Community Guidelines and the Terms of Service, which broadly prohibit use of Snapchat for illegal activities in the user's jurisdiction and expressly prohibit categories of content illegal in the EU, including Hate Speech and content advocating terrorism or violent extremism.

EU users can submit an Article 16 notice via the in-app or web reporting flow by selecting the illegal content reporting option. Reports are first assessed for compliance with the Community Guidelines. [redacted]

[redacted]

[redacted]

[128] Snap Inc., Response to the European Commission's Request for Information, 25 June 2026, response to Question 34.

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

As part of its continuous risk assessment process, Snap monitors processing times for notices submitted through the dedicated illegal-content channel. [redacted]

Protections against misuse (Article 23 DSA)

Snap maintains processes to address misuse of the moderation infrastructure, including: [redacted]

[129] Snap Inc., Response to the European Commission's Request for Information, 25 June 2026, response addressing Article 16 notice-processing times.

[redacted] The policy on misuse and the factors considered are set out in the Terms.

5.4.7 Moderation Effectiveness Metrics

Consistent with the methodology refinement in Section 3.6 on effectiveness metrics disclosure, and with the disclosure requirements in Articles 15(1) and 24(1) DSA, the following metrics are used to assess moderation effectiveness and are disclosed where available for the 2026 assessment period:

Metric	2026 disclosure
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
Notice volumes (Article 16)	[redacted]
Internal complaint volumes (Article 20)	[redacted]

[redacted]

Out-of-court dispute settlement (Article 21)	
Suspensions for misuse (Article 23)	

Prevalence Testing methodology, oversight and results are addressed in detail in Section 6. Transparency Report publication and cadence are addressed in Section 5.3 (Transparency) and Section 6.

5.4.8 Conclusion

The content moderation mitigations described in this Section operate across the proactive detection suite, the Search and contact surface measures, the product-specific moderation workflows, and the reactive reporting pathways. In combination with the design mitigations in Section 5.1, the Terms framework in Section 5.2, the transparency measures in Section 5.3, the enforcement framework in Section 5.5, and the other sub-sections of Section 5, they support the Article 35(1) proportionality analysis recorded in the corresponding risk assessments in Section 4.

5.5 Enforcement

Introduction

Section 5.5 sets out Snap's policies and processes for the enforcement of the Snap Terms (Section 5.2) in respect of Snapchat's in-scope services. Enforcement is integral to the risk-mitigation framework recorded across Section 4 of this Report, the substantive Terms identify the prohibited categories of content and conduct, and the enforcement framework operates the policies and processes through which those Terms are applied. The framework operates with the objective of protecting Snapchatters and the broader communities served by Snapchat, in a manner that is transparent, consistent and equitable, and that balances safety considerations with the privacy interests of users.

[redacted]

Increases over time in reporting volumes, enforcement actions and proactive law-enforcement referrals do not, as a general matter, indicate that Snapchat has become less safe. The recent reporting periods have seen these upward trends in operational activity correlate with a continued reduction in Policy Violating Prevalence (PVP) on Snapchat overall, recorded in detail in Section 6.4 (Prevalence Testing) and across the substantive Category 1 assessments of Section 4. The framework operates on the basis that as detection and enforcement against violations expand in scope and capability, the underlying frequency of violations on Snapchat continues to decrease.

[redacted]

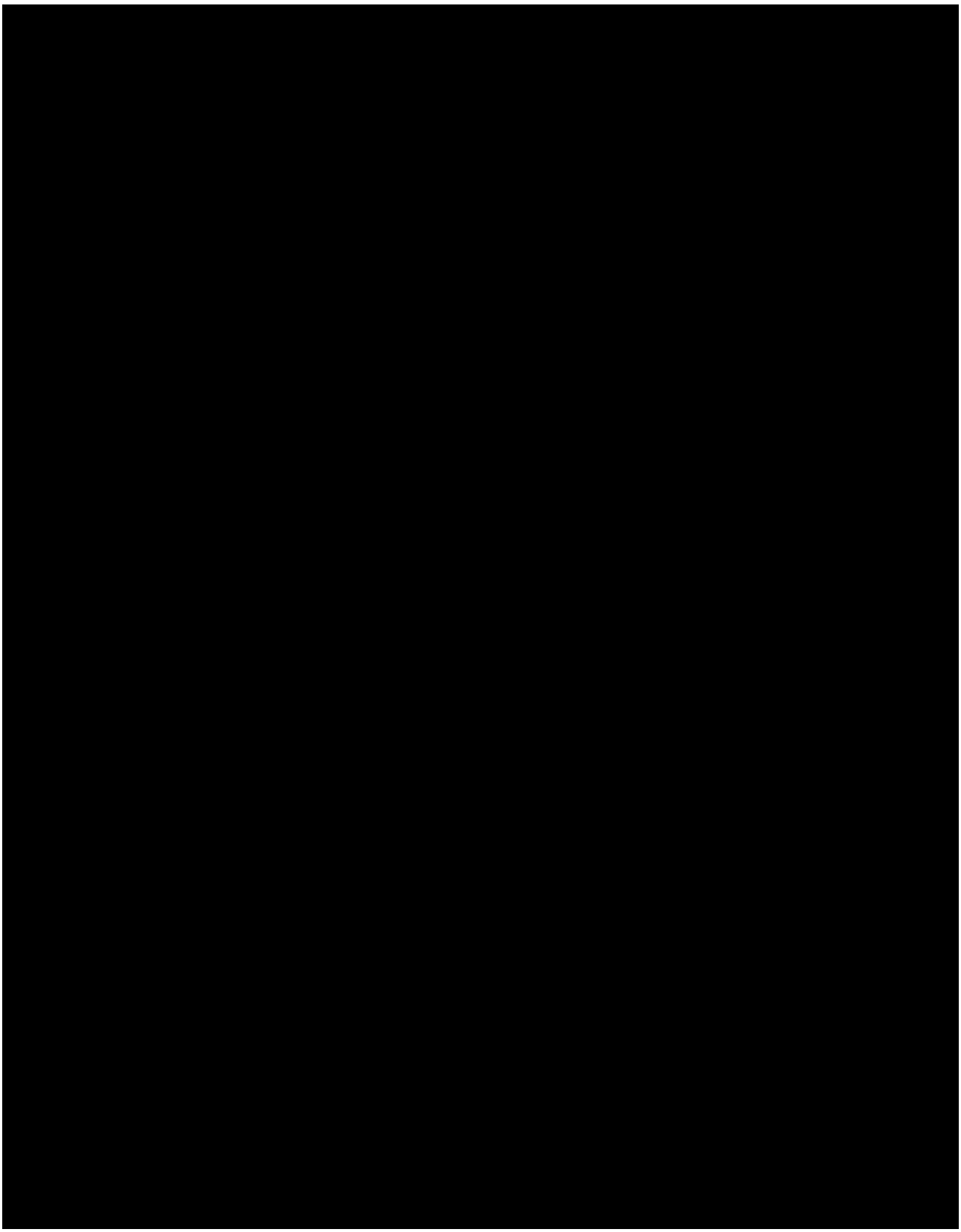
Review and Enforcement

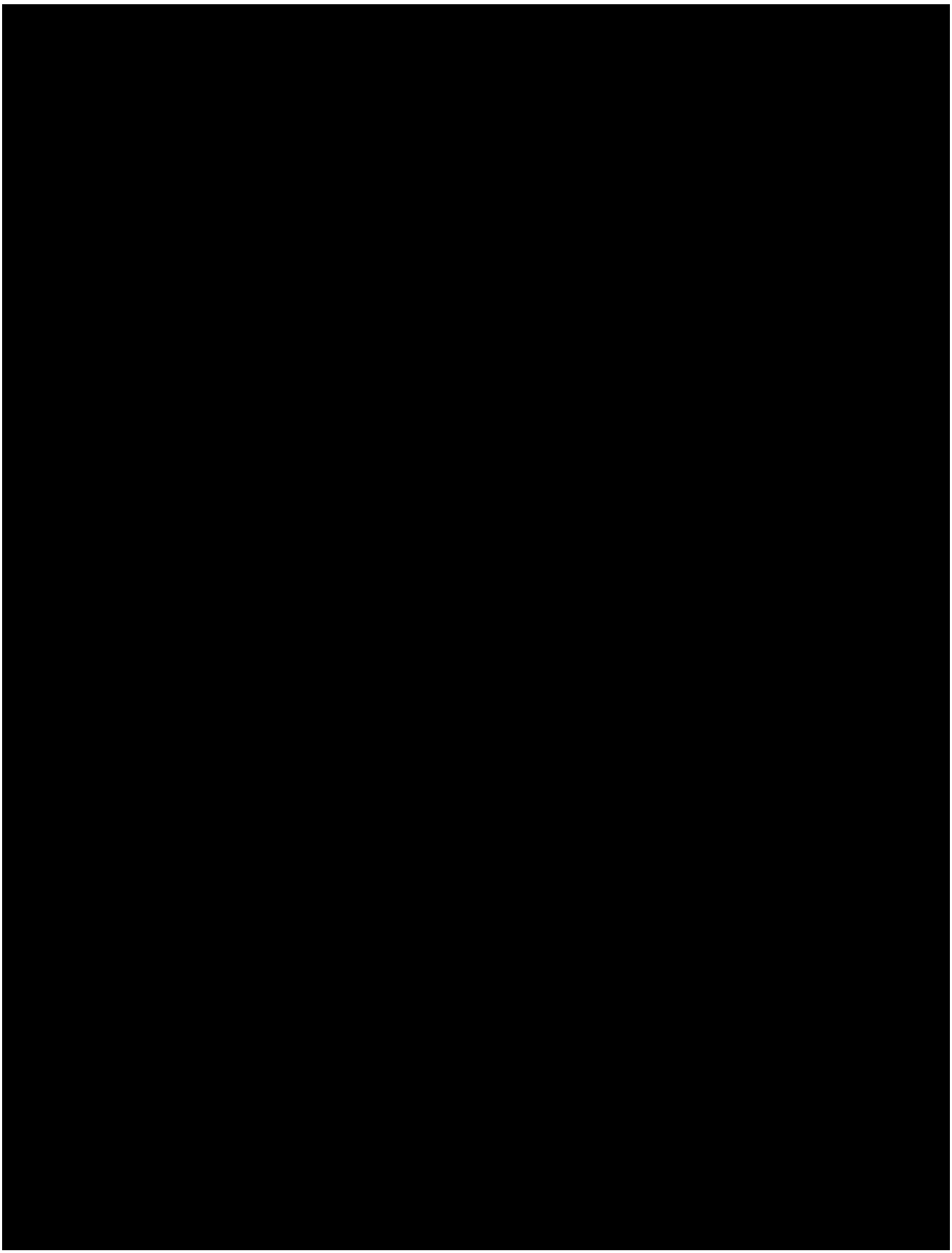
All flagged or reported accounts and content are reviewed against the Snap Terms. The Community Guidelines and the Terms of Service are applied when determining whether to take enforcement action against an account. The Community Guidelines, the Terms of Service, and the Content Guidelines for Recommendation Eligibility are applied to moderate Snaps posted to Discover (the Stories tab within the Discover section) and Spotlight.

Enforcement types

[redacted]

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]





[redacted]

Transparency

Snap publishes information regarding the enforcement framework and the misuse policies through several public-facing instruments:

- **Community Guidelines.** Including the Severe Harms Explainer and the associated Harm Category Explainers linked from the Community Guidelines.
- **Snapchat Moderation, Enforcement, and Appeals Explainer.** Explains the content moderation practices, enforcement, and appeals processes, including a short guide to the strike system. Detailed strike thresholds and related timeframes are not disclosed publicly: these parameters are subject to evaluation and may change as part of internal review processes, and public disclosure could enable users to game the enforcement system.
- **Content Guidelines for Recommendation Eligibility.** Set additional standards beyond the Community Guidelines for content eligible for algorithmic distribution. The Content Guidelines apply standards calibrated to a child-population baseline (with the operational assumption that users as young as 13 may be viewing content) and limit the algorithmic

[redacted]

distribution of content that, while not illegal or harmful, may be unsuitable in certain circumstances.

Statement-of-reasons transparency under Article 17 DSA is addressed through the notice mechanisms recorded in the Notice and Appeals System sub-section below. Public-facing transparency on enforcement metrics is provided through Snap's Transparency Reports, recorded in Section 5.3.

Notification of Criminal Offences (Article 18 DSA)

Proactive referrals to law enforcement and governmental agencies

Snap engages with law enforcement and governmental agencies to support safety on the platform. [redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

Law-enforcement takedown orders (Article 9 DSA)

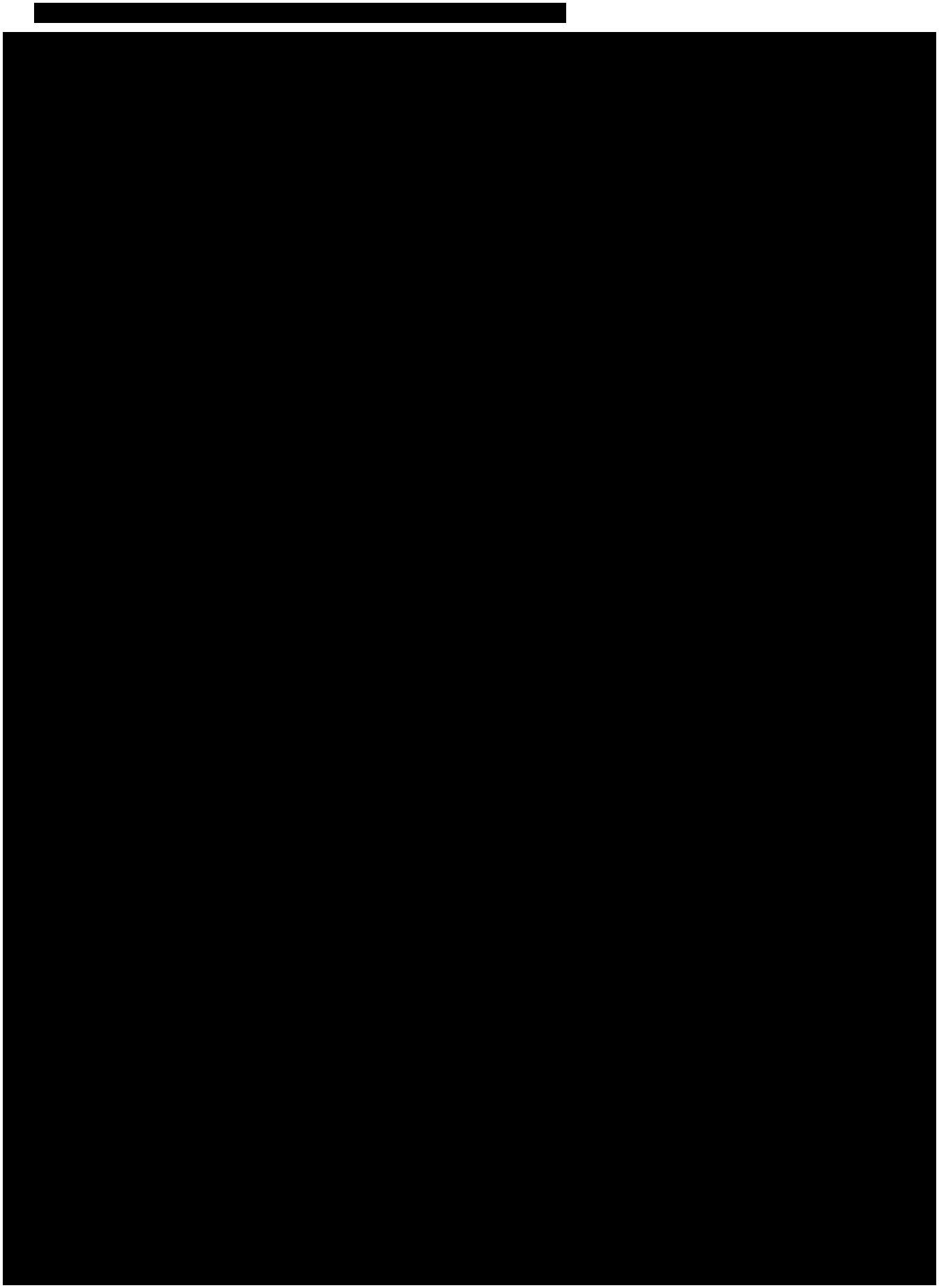
[redacted]

- [redacted]
- [redacted]

Email-submitted orders

[redacted]

[redacted]



[redacted]

Notice and Appeals System (Articles 17, 20, 21 DSA)

Snap's notice and appeals framework operates to ensure that Snapchatters have a clear understanding of why an action has been taken against their account or content (Article 17 DSA), and to provide a meaningful pathway to dispute enforcement outcomes (Article 20 DSA). The framework supports the underlying objectives of safeguarding the interests of the user community while preserving Snapchatters' rights.

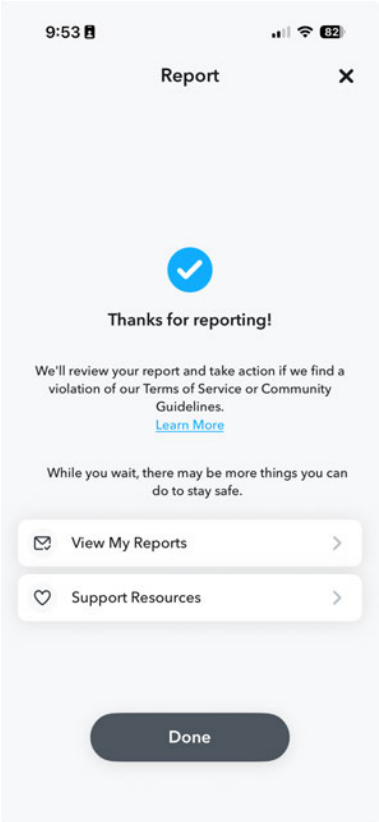
[redacted]

[redacted]

Notice and appeals eligibility depends on the relevant policy and the type of content violation. The substantive operation of the notice and appeals framework is described below.

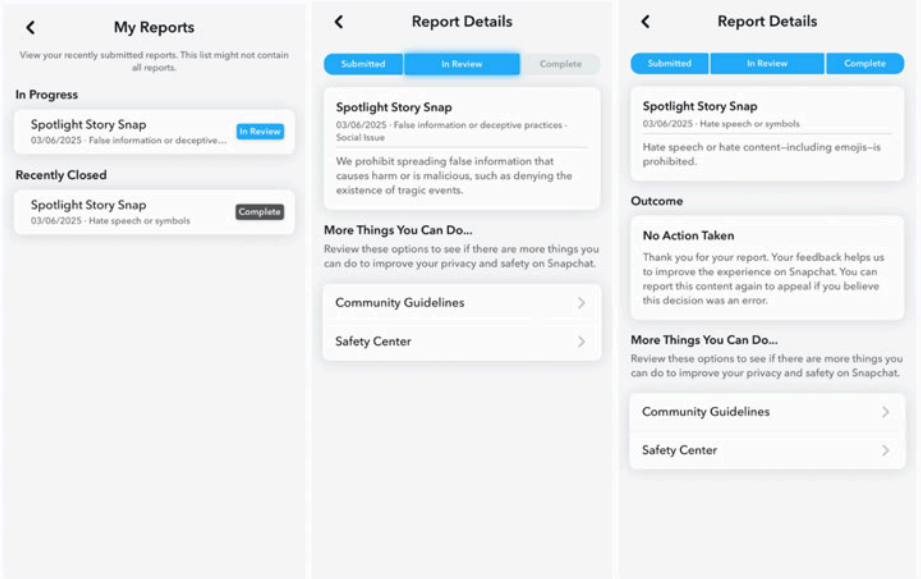
Notice to Reporter

Where a reporter submits a report in-app concerning content or an account possibly violating the Terms or otherwise causing harm, Snap confirms receipt of the report and provides assurance that the matter is under investigation.



Following completion of review, the reporter is sent an update on the outcome of the review, in-app and/or by email. Snapchat users who submit a report in-app can review the status of their reports in the My Reports dashboard, accessible from within the app via My Account > Settings > My Reports.

[redacted]

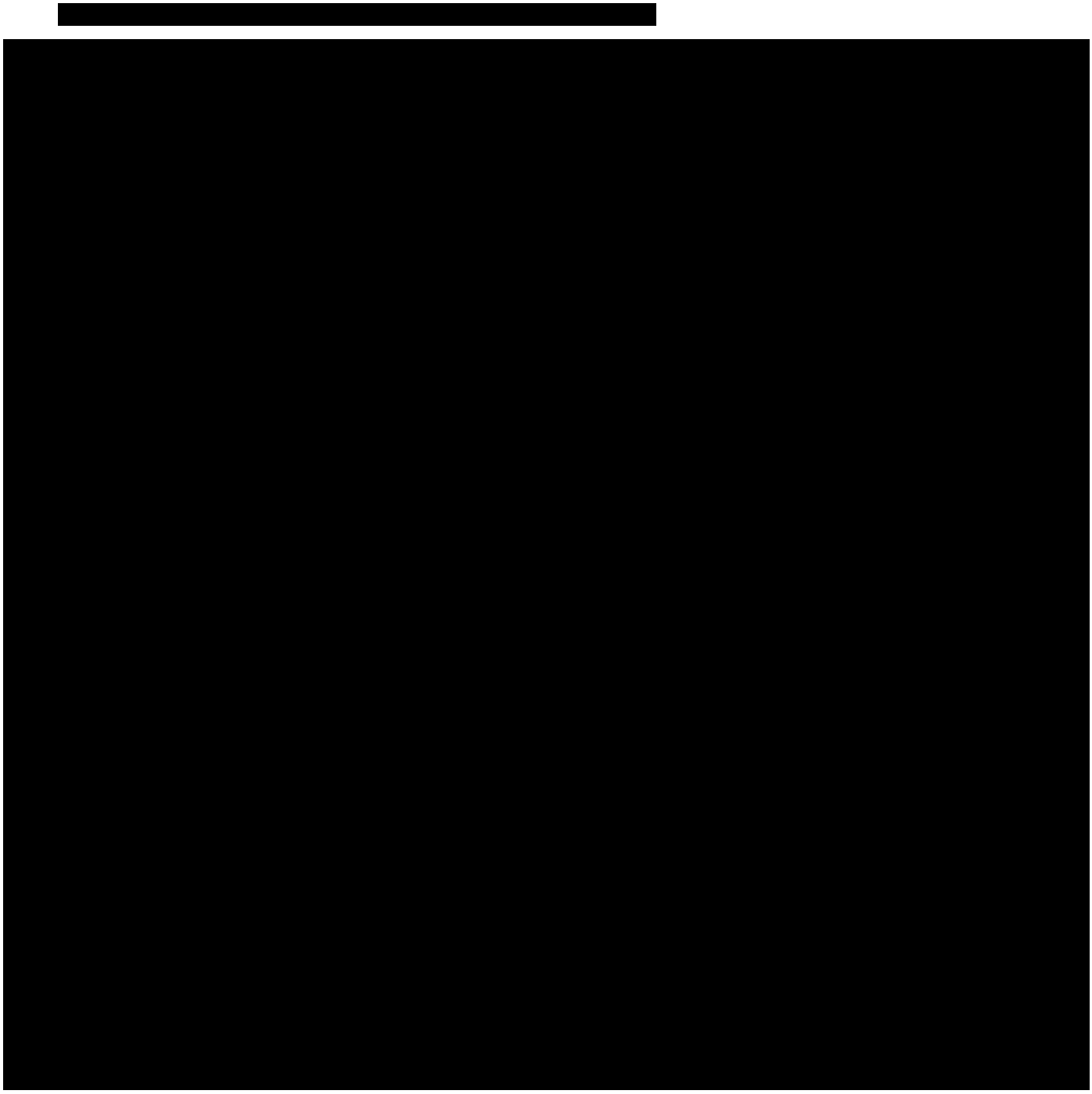


[redacted]

Account-Level Notice and Appeals

Where Snap takes enforcement action against a user's content or account, whether following a report or proactive detection, Snap sends a notification to the user indicating that action has been taken (a warning or a lock) and the reason for the action. [redacted]





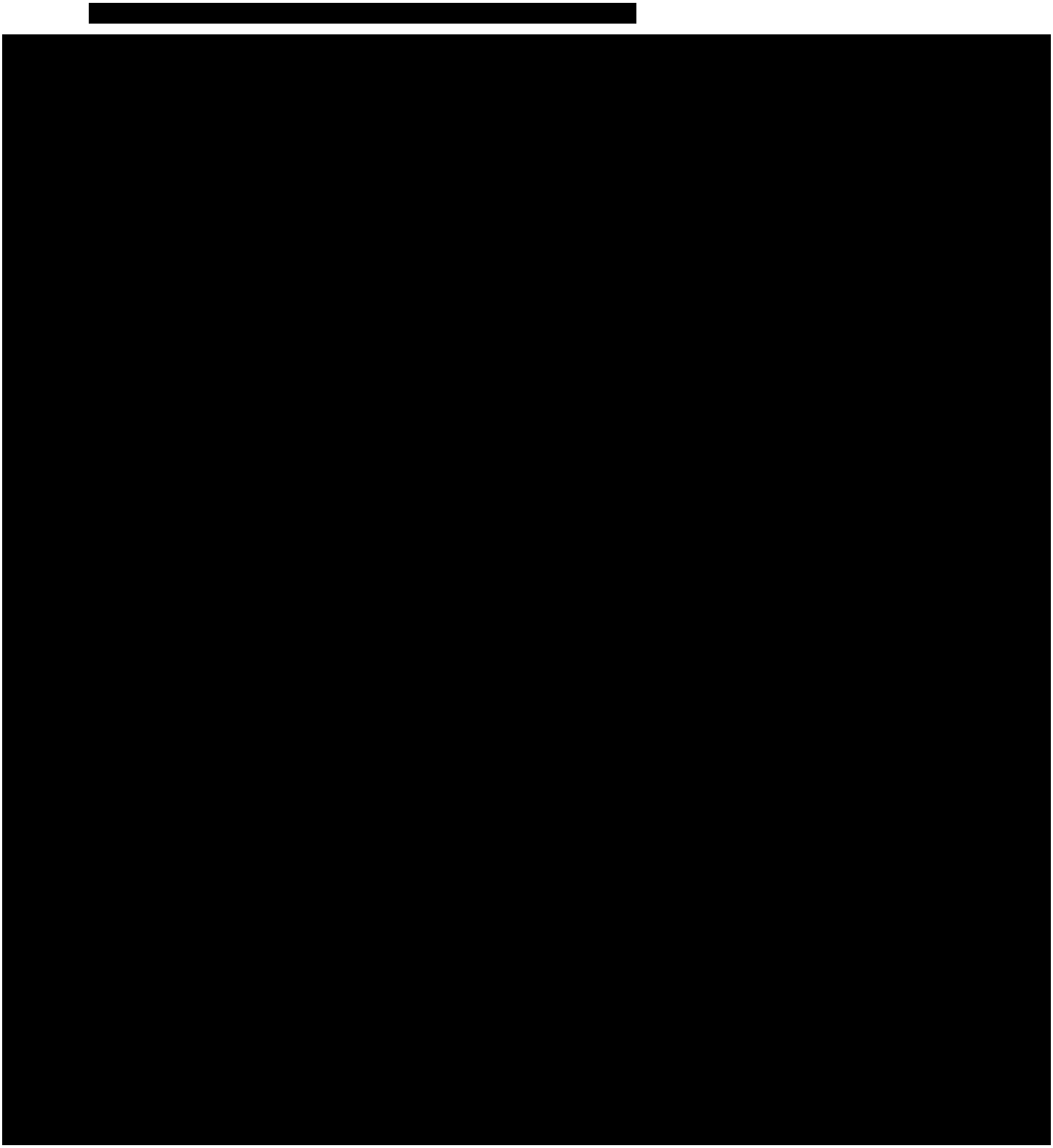
[redacted]



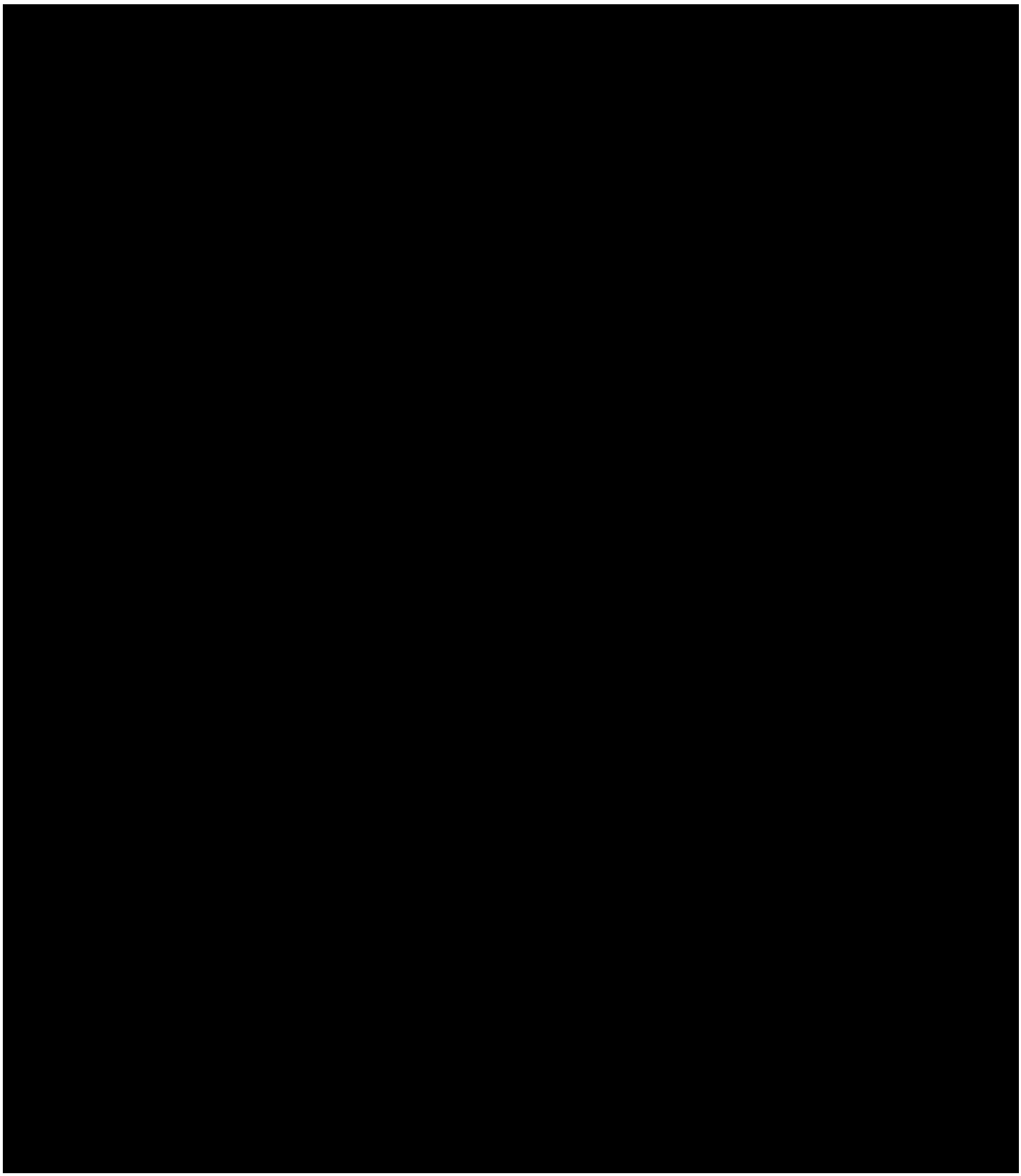
Upon submission of the appeal, an overview of the appeal process is provided to the user. In most cases, the user is permitted to download their data while the appeal is under review.

Appeal outcome notification

Snap notifies users of the outcome of their appeal in a timely manner. Where Snap grants an appeal of an account lock, access to the Snapchatter's account is restored. Where Snap denies an appeal, the user is invited to download their data where eligible to do so.



[redacted]



Content-Level Notice and Appeals





Content appeal request form

The content appeal request form prompts the content creator to enter the Appeal ID from the enforcement notice and any additional information requested as part of the appeal process.

[redacted]



Content appeal outcome

Where the content appeal is granted (the original content moderation decision is overturned), the content is reinstated and the content creator receives a notice confirming the outcome.

[redacted]



Where the content appeal is rejected, the original content moderation decision remains in effect and the content creator is informed of the outcome.

[redacted]

[redacted]

Effectiveness of Enforcement

The performance of Snap's reporting and enforcement systems is recorded in Snap's Transparency Reports, with country-level breakdown for EU Member States. The substantive Category 1 assessments of Section 4 record the category-specific enforcement metrics (Article 17 statement-of-reasons volumes, median turnaround times, post-publication report volumes, and similar data) for the relevant assessment cycles.

Stakeholder engagement

Snap engages with NGOs and other stakeholders on a continuing basis to discuss the operation of the enforcement framework. [redacted]

[redacted]

Continuous improvement and prevalence trajectory

Prevention, detection and removal of illegal and other violating content remain ongoing operational priorities, with continuing analysis of enforced content and accounts to support the development of proactive measures targeting bad-actor patterns. The principal Category 1 risks assessed in this Report remain in the lowest likelihood category, and H1 2026 prevalence results remain low across the relevant harm categories, notwithstanding limited variation in individual categories.

[redacted]

[redacted]

Conclusion

The enforcement framework recorded in this Section, in combination with the other mitigations recorded across Section 5, supports the conclusions recorded across the Category 1, Category 2, Category 3 and Category 4 assessments of Section 4 that Snap's mitigations for the systemic risks identified in respect of Snapchat's in-scope services are reasonable, proportionate and effective. The framework operates through the integrated architecture described above, with operational performance monitored on a continuing basis through the Section 6 risk-detection processes.

Increases over time in reporting volumes, enforcement actions and proactive law-enforcement referrals correlate with a continued reduction in Policy Violating Prevalence on Snapchat overall - supporting the determination that the framework is operating as intended. Continuous improvement of the enforcement framework will operate through the Section 6 risk-detection processes during the 2026 assessment period and beyond, with particular attention to the operational outputs of the 2026 reporting-pathway changes recorded in Section 5.4.6, the operational outputs of the minor-protection refinements recorded across Section 5.1.3, Section 5.4.4 and Section 5.8, and the continued evolution of cross-platform signal-sharing arrangements supporting the framework.

5.6 Algorithmic Systems

Introduction

Section 5.6 sets out the specific mitigation measures that Snap has put in place in respect of Snapchat's algorithmic systems for the in-scope services to address the risks identified in the risk assessment recorded in Section 4. Algorithmic content recommendation systems present a number of risks that are mitigated through the integrated framework recorded across Section 5: substantive Terms (Section 5.2), automated and human moderation (Section 5.4), enforcement (Section 5.5), public-facing transparency (Section 5.3), the personalisation opt-out under Article 38 DSA, and the additional mitigations recorded below.

Consistent with Article 34(1) DSA, Snap's risk assessment is proportionate to the risks identified, taking into account their severity and probability and the design of the recommender systems and other relevant algorithmic systems. While Snapchat operates several algorithmic systems across the in-scope services (Spotlight, Discover, Snap Map, Lenses, Public Profiles, and Advertising), the risks identified in the Section 4 risk assessment focus on the following systems:

- Content Recommendation Systems in Spotlight and Discover;
- Advertising Systems;
- Content Moderation Systems.

The specific mitigations for the Content Moderation Systems and the Advertising Systems are recorded in Sections 5.4 and 5.7 respectively. This Section 5.6 is principally focused on the specific mitigations for the Content Recommendation Systems in Spotlight, Discover, Snap Map and Lenses (referred to in this Section as the "content recommender systems").

In-Scope Algorithmic Systems

Spotlight

Spotlight is Snap's short-form video feed located primarily on the fifth tab of the Snapchat app. Spotlight provides users with a way to view content created and submitted by the Snapchat community via a personalised feed. Users can post videos to Spotlight either via the Snapchat app or the website; videos on Spotlight are public and visible to users on the Snapchat app and on the web, and a link to the Spotlight video can be shared to other platforms. Viewers can indicate dislike ("I do not like…") for a Spotlight creator or content, or can indicate fondness by selecting a heart icon.

Spotlight content is personalised. Spotlight's ranking algorithm is described in Snap's public-facing documentation[130]. Users may opt out of personalisation as recorded in the same documentation.

[130] How We Rank Content on Spotlight.

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

Discover

Discover is part of the fourth tab in the Snapchat app and is dedicated to Creator Stories, including Media Partner content and certain user-generated content ("UGC") created from Snaps by popular users ("Creator Content"). The UGC that appears on Discover includes the Public Stories from Snap Stars and other users meeting a follower count threshold.

Discover displays personalised content using its ranking algorithm, described in Snap's public-facing documentation[131]. The intended purpose of the processing is to personalise Discover and to support discovery of new content relevant to the user's interests. Users may opt out of personalisation as recorded in the same documentation.

[redacted]

[131] How We Rank Content on Discover.

[redacted]

Snap Map

Through an interactive experience, Snap Map shows users what is happening nearby and around the world, anchored by the context of friends' Bitmojis. Snap Map starts with the user at the center and reflects the people, places and events the user cares about, and supports finding friends, expression, finding things to do, and exploring places elsewhere.

Eligible Snapchatters can share their Snaps to the Map by selecting "Snap Map" on the Send To page. Where the Snapchatter has a Public Profile or is sharing the My Story with everyone, the Snap may also be shared on Snap Map when it is tagged to a place or venue. Snap Map features five types of user-generated content: Map Stories, Place Stories, City Stories, the Heat Map, and Friend Stories tagged with places. [redacted]

[redacted]

Lenses

Snapchat Lenses are Augmented Reality (AR) experiences designed to transform the way users look (front-camera experiences) and the world around them (rear-camera experiences). Lenses are used principally for entertainment purposes, including the creation of Snaps with 3D effects, digital objects, characters, and transformations to image and voice. Snapchatters can interact with Lenses in the Lens Carousel, via Search, and via Lens Explorer. Lenses are created by

[132] How We Rank Content on Maps.

[redacted]

community developers, advertisers, and Snap's internal Lens Team. [redacted]

[redacted]

[redacted].

Public Profiles

Public Profiles enable Snapchatters to be discovered and followed in the app and to showcase their public Snaps, Lenses, Stories, Spotlights and other information. [redacted] Snapchatters with a Public Profile that are particularly active may have their accounts upgraded to a Creator Account; Creator Accounts are eligible to have their content shown in the Discover section of Snapchat.

[redacted]

Advertising

Snap relies on online advertising to support its business. Snap Advertising is a digital ad product that allows advertisers to create and manage ads (images, videos and lenses) that target relevant audiences on Snapchat. [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Content Moderation

Content moderation operates through a series of proactive content scanning and responsive content moderation procedures. [redacted]

[redacted]

[redacted]

Adaptation of Algorithmic Systems to Mitigate Systemic Risk

Enabling user choice in content prioritisation (Article 38 DSA)

Snapchatters in the European Union have the option to disable personalised public content recommendations. In Discover and Spotlight, users can disable personalised content by either: (i) selecting "..." on a piece of content and then "Why am I seeing this content?", which directs the user to Settings; or (ii) navigating directly to Settings and "European Union Controls".

[redacted]

Where users disable personalisation, public content is recommended on the basis of basic data only (for example, the language set on the phone, age, and country). Users continue to see content but the content is more random and less tailored to interests. Users can also modify ad content prioritisation preferences across the platform by modifying notifications received by type or topic, toggling Lifestyle Categories, and adjusting ad preferences within Settings.

Oversight and Administration

Roles and responsibilities



Algorithmic System Review

[redacted]

[redacted]

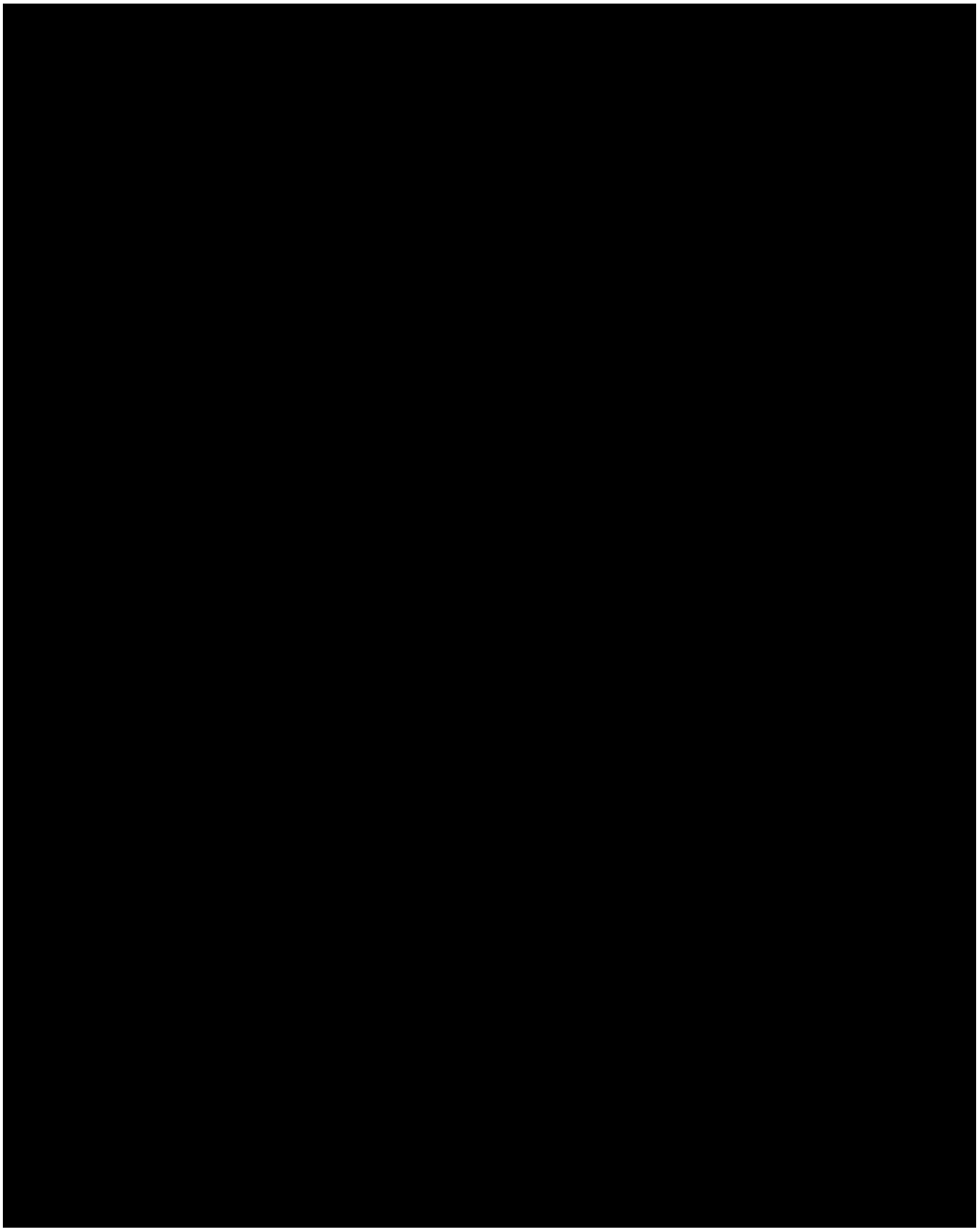
[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]





[redacted]

[redacted]

Pre-launch testing

Snap conducts model offline testing where appropriate. Offline testing supports quick, iterative model performance adjustments based on standard training sets. [redacted]

[redacted]

Pre-launch legal review

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

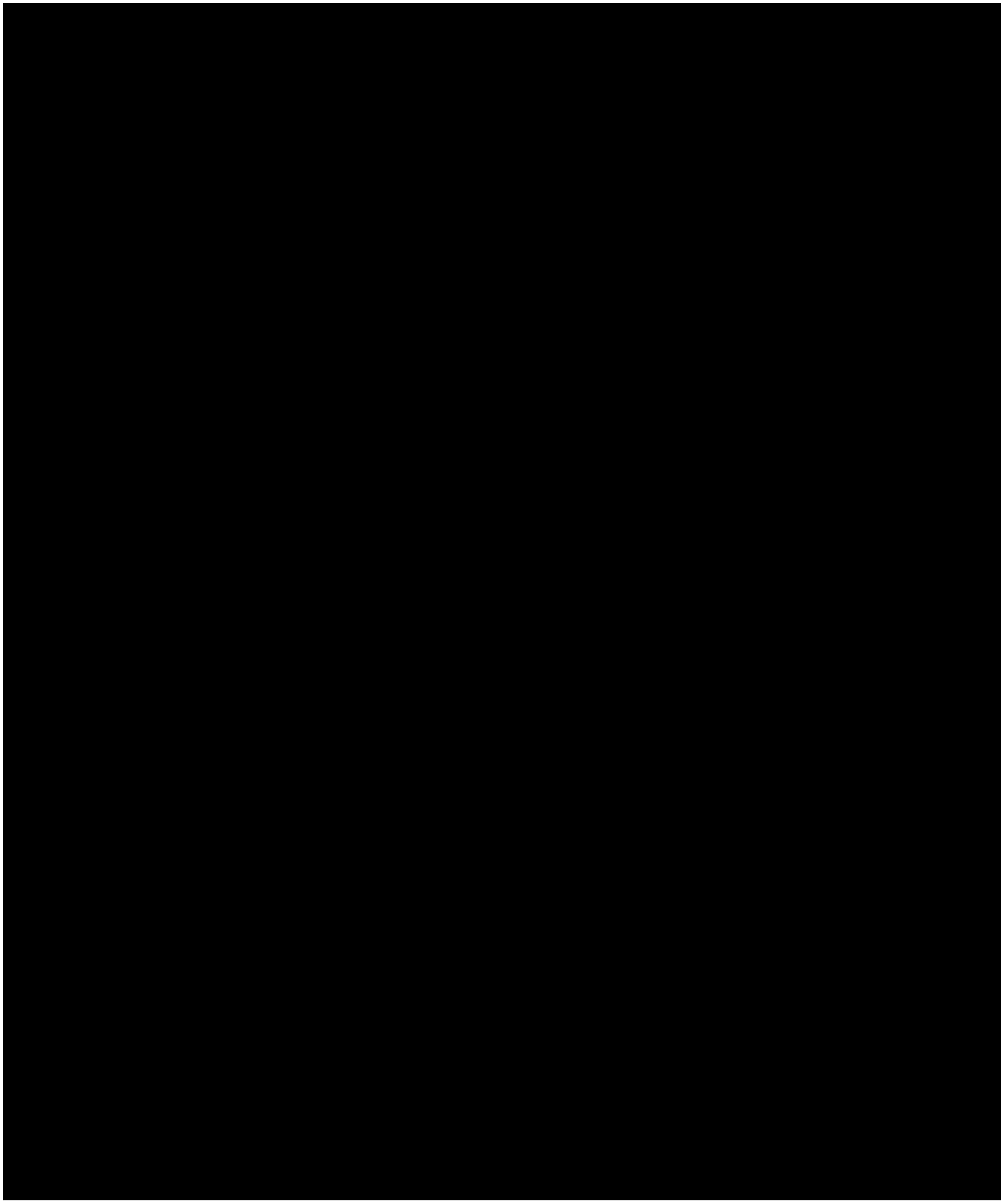
- [redacted]

Adaptation and Testing (Article 35(1)(d) DSA)

Consistent with Article 35(1)(d) DSA, this sub-section records the adaptation and testing of the algorithmic recommendation systems to address the risks identified in Section 4 of this Report. The adaptation framework operates against an integrated set of risks; the table below summarises: (i) the specific risks identified that are relevant to the algorithmic recommendation systems; (ii) the adaptations applied to mitigate those risks; and (iii) the evaluation methods through which mitigation effectiveness is monitored.

[redacted]

[redacted]



Illegal or violating content

As outlined in Section 5.2 (Terms), all content on Snap is required to comply with the Terms (which require all public content on Snapchat to be suitable for users as young as 13), including the Community Guidelines. Content personalised through the algorithmic recommendation system is required to comply with the more restrictive Content Guidelines for Recommendation Eligibility.

As detailed in Section 5.4 (Moderation), the recommender systems and processes are adapted to enforce the content policies through automated and human moderation.

As detailed in Section 5.5 (Enforcement), users may report inappropriate and illegal content. Each piece of content in Spotlight and Discover has a menu allowing users to report content. Where content is determined to violate the policies , it is made ineligible for recommendation by the algorithmic systems.

Effectiveness of these measures is tested through Prevalence Testing (Section 6.4) and through review of privacy and other consumer queries received by the community support teams, the Data Protection Officer, and the DSA Compliance Function.

Lack of user understanding

The recommender systems are complex; the process, the signals used in ranking, and the relative significance of each signal may be challenging for users to understand. To support user understanding and to address frequently asked questions, and as part of DSA compliance, Snap has adapted its public-facing content to include articles describing how content is personalised in Spotlight, Discover, Maps, and Ads, including a description of the main parameters used for the recommender systems and the weighting applied to each signal. Users may also reach out to the Support team with concerns or questions about how the algorithms work. The framework is reviewed through assessment of privacy and consumer queries received by the community support teams and the Data Protection Officer.

Intrusive personalised recommendations

Personalisation is intended to support content relevance for the user, but there is a risk that some users may experience personalised recommendations as intrusive. The Article 38 DSA opt-out framework recorded above operates against this risk: users in Discover and Spotlight can disable personalised content through European Union Controls in Settings; users can re-enable personalisation by selecting the favourite icon in Discover or Spotlight and selecting Enable, or by navigating directly to Settings and European Union Controls. Users can disable personalised ads and Lenses through Personalisation Controls in Settings; where disabled, content ads, and lenses are recommended on the basis of basic data only (country and age) rather than user profile. The Reset Suggested Content feature (accessible in Settings) supports users in refreshing their public content feeds. When users enable Reset Suggested Content, most indicators used for

[redacted]

personalisation of public content feeds are reset (though not all): users may continue to see content or creators they have followed and content tailored to their demographic, and past blocks and dislikes continue to be taken into account.

Discrimination

Algorithms processing special categories of personal data (as defined in the GDPR) on a large scale are recognised as high-risk and require explicit user consent. The recommender systems are adapted not to tag or identify special categories of personal data, including for the purpose of recommending content and ads. Cross-reference to Section 4.2.5 (Right to Non-Discrimination).

Rapid and widespread illegal or false content; crisis exposure

There is a risk of rapid and widespread distribution of illegal or false content on Spotlight, Discover and Map, and of exposure to crisis situations and unexpected events (for example, riots). [redacted]

Filter bubbles

The recommender system algorithms are designed to serve users with content the user is likely to find engaging, based in part on categories of content the user has previously watched. Without safeguards, this design pattern presents a risk that the algorithm tags users who view content that may not be harmful on its own as being interested in that content, with repeated and frequent exposure potentially producing harm. By way of example, while one piece of content related to dieting may not be harmful, repeated and frequent diet content exposure may produce inappropriate pressure or a skewed perspective on relationships with food.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Erroneous content exclusion (over-moderation)

There is a risk that efforts to ensure appropriate content on Snapchat result in some appropriate content being mistakenly identified and incorrectly moderated. The risk engages users' rights to freedom of expression (cross-reference to Section 4.2.2). [redacted]

[redacted]

[redacted]

[redacted]

[

[redacted]

[redacted]

[redacted]

[redacted]

Quality assurance signals

Snap monitors algorithmic systems related to content moderation for quality and precision on a continuing basis. Monitoring takes into account the following signals:

Signal	Description
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Snap uses[redacted]

[redacted]

Conclusion

The algorithmic-systems framework recorded in this Section, in combination with the other mitigations recorded across Section 5, supports the conclusions recorded across the Category 1, Category 2, Category 3 and Category 4 risk assessments of Section 4 that Snap's mitigations for the systemic risks identified in respect of Snapchat's in-scope services are reasonable, proportionate and effective. The framework operates through the integrated architecture described above: substantive Terms; Content Guidelines for Recommendation Eligibility calibrated to a child-population baseline; risk-based moderation across Spotlight and Discover; user-facing transparency on recommender system operation; the Article 38 DSA personalisation opt-out through European Union Controls; the Article 25 DSA dark-pattern review framework; the Article 27 DSA recommender systems transparency; and the integrated PASS Review and DDIA framework operationalising privacy by design.

5.7 Advertising Systems

Introduction

Snap relies on online advertising as the principal source of revenue supporting the operation of Snapchat's services. Advertising systems are recognised as having a significant capacity to give rise to the systemic-risk concerns referenced in Article 34 DSA: where not adapted and tested appropriately, advertising systems can give rise to, amplify or result in the rapid and wide dissemination of illegal content and other harms identified across Section 4 of this Report.

Snapchat Ads Manager is the principal product through which advertisers create and manage advertising campaigns targeting relevant audiences on Snapchat. Snap processes information about Snapchatters to deliver advertising within Snapchat. Section 5.7 sets out how Snapchat's advertising systems work, the benefits the framework provides, and the targeted measures adopted in line with Article 35(1)(e) DSA to mitigate the risks identified in Section 4 of this Report.

This Section is cross-referenced extensively across the Section 4 risk assessments, in particular Section 4.1.3 (Sale of Prohibited Products), Section 4.1.10 (Harmful False Information), Section 4.1.11 (Fraud and Spam), Section 4.2.4 (Right to Data Protection), Section 4.2.5 (Right to Non-Discrimination), Section 4.2.7 (Right to Consumer Protection), Section 4.3.1 (Negative Effects on Democratic and Electoral Processes), and Section 4.4.3 (Negative Effects on Minors).

How Advertising Systems Work

An overview of Snap's advertising services is publicly available through Snap's Business Help Center[133]. Snap collects data about Snapchatters as they register, log in and use Snapchat. The Privacy Policy describes the categories of data processed in connection with advertising:

- information that the user provides to Snap;

[133] https://businesshelp.snapchat.com/s/?language=en_US

[redacted]

- information that Snap collects as the user interacts with Snapchat;
- information that Snap collects from third parties[134].

Advertiser-side tools

Snapchat Ads Manager and its associated tools allow advertisers to apply this data to advertising campaigns. Advertisers can use the Audience Insights tools to view estimated aggregated demographics (including age), locations, interests and device overviews of their targeted audience. User-level data is not directly available to advertisers through these dashboards. Several advertising tools provide controlled mechanisms for advertisers to use data about their own customers (subject to data-protection-aligned constraints):

- **Snap Custom List Audiences.** Allows an advertiser (or its agent) to upload customer list data via Ads Manager. The customer list data is used to create an "audience" of Snapchatters matching the information in the list, supporting targeting to that audience or to similar audiences on Snapchat.
- **Snap Pixel and Conversion API.** Pixel is a JavaScript code installed on advertiser web pages that sends data to Snap when those pages are accessed. Conversion API is server-to-server code installed on advertiser servers that passes web, app and offline events directly to Snap.
- **Advanced and Estimated Conversion.** Additional services that support advertiser targeting and performance measurement using privacy-enhancing techniques.

Snap acts as a data processor of data relating to EU data subjects received from advertisers via the [redacted] processing the information in accordance with advertiser instructions and subject to a data processing agreement that follows the requirements of Article 28 GDPR[135].

Ad ranking algorithm

The ad ranking algorithm determines which advertisements are presented to a Snapchatter who falls within the selected audience for those advertisements. [redacted]

[134] Snap Inc., 'Snapchat Ads Transparency'; Snap Inc., 'Snap Advertiser Privacy and Security FAQs' (accessed 4 August 2026).
[135] Snap Inc., 'Snap Advertiser Privacy and Security FAQs', section 'Terms and Conditions'; Snap Inc., 'Data Processing Agreement' (versions applicable during the 2026 assessment period).

[redacted]

[redacted]

Benefits

Snapchat is used by a substantial population in the European Union, where it supports user interaction, communication and access to public content. Targeted advertising is the principal mechanism through which Snap is able to maintain and develop Snapchat as a service available without subscription fee. The framework supports several distinct beneficiary groups:

- **Snapchatters generally.** Snapchat is maintained and developed for the benefit of all recipients regardless of their ability to pay. A subscription-fee-only service would restrict access to Snapchat - with particular implications for fundamental EU rights to information and access to services, including in respect of users in the 13–17 cohort.
- **Minor users specifically.** Access to Snapchat without a subscription fee supports the exercise by 13–17-year-olds of their digital rights (as recognised under Article 24 EU Charter and the UNCRC), including access to information, association with others, voice, play and recreation, regardless of their financial background. Substantive treatment recorded in Section 4.2.6 (Children's Rights) and Section 4.4.3 (Negative Effects on Minors).
- **Snapchatters in respect of advertising relevance.** Targeted advertising supports the delivery of advertisements that are more relevant to user interests, age and context. Snapchatters are provided with controls to manually hide advertisements and to manage data preferences (recorded below).
- **Advertisers.** Snapchat Ads Manager supports advertisers in promoting brand and product information to relevant audiences and in measuring campaign performance. In a measurement study conducted under a partnership announced in March 2024, Spotlight and Creator content on Snapchat was assessed as 99% brand safe during the studied period. Snap also launched a brand-safety reporting solution in June 2024 to provide advertisers with information regarding the content against which their advertisements appeared.[136]

The framework operates on the basis that targeted advertising supports the legitimate benefits described above only where the processing of personal data (including by way of profiling) does not result in users being subject to economic exploitation. Privacy and safety considerations have been integrated into the architectural design of the advertising systems from initial deployment, and the framework has been adapted on a continuing basis to mitigate the risks identified across Section 4 of this Report.

[136] Snap Inc., 'Snap Partners With Integral Ad Science, Announces New Brand Safety Solutions' (26 March 2024); Snap Inc., 'Snap, IAS Brand Safety & Suitability Measurement Solution Now Available' (17 June 2024).

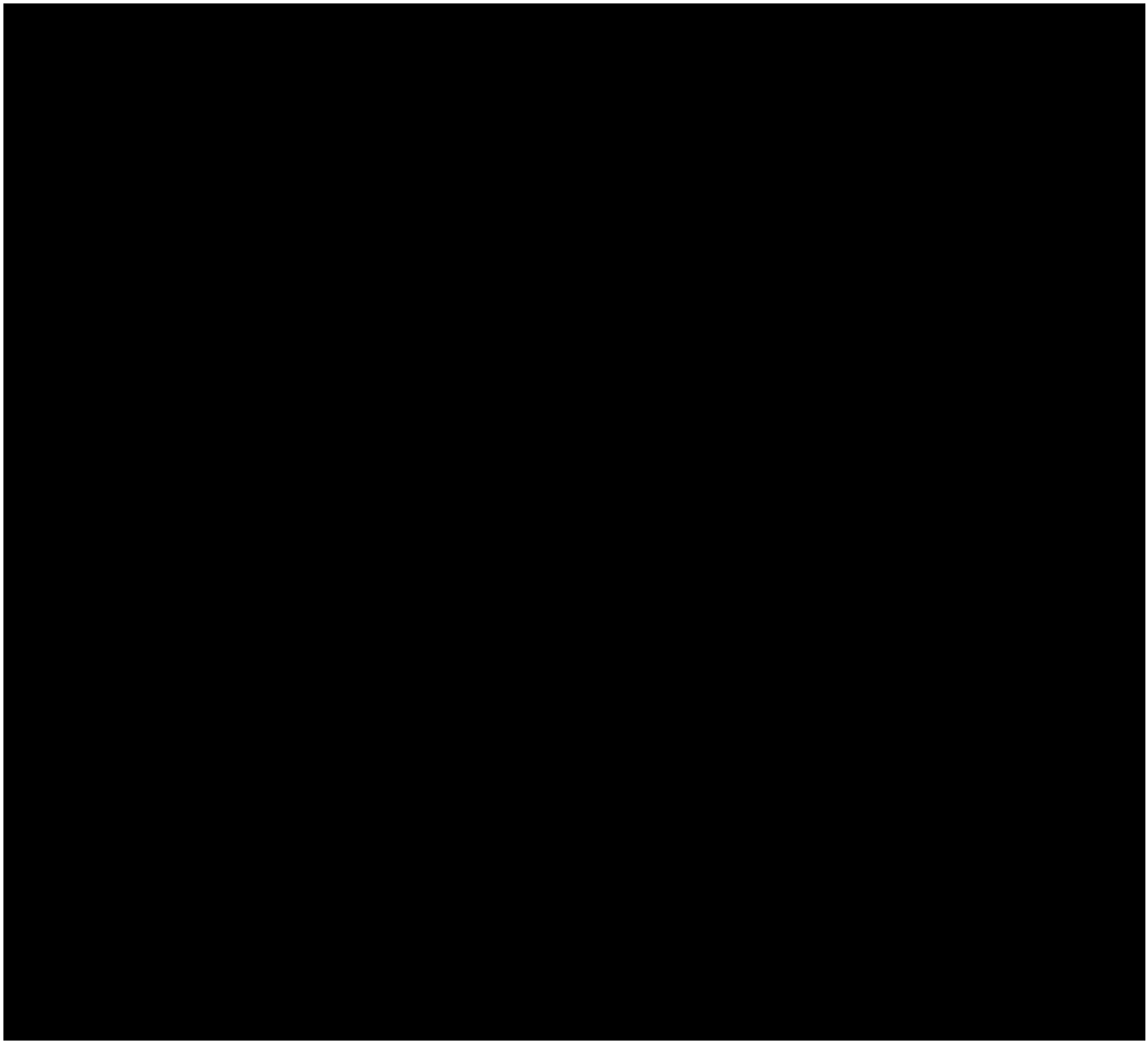
[redacted]

Adaptation and Testing

Consistent with Article 35(1)(e) DSA, Snap has adapted Snapchat's advertising systems and adopted targeted measures aimed at mitigating the risks identified in Section 4. The framework operates through the integrated set of design, policy, review, transparency and user-control measures described below.

Summary of risks, mitigations and evaluation methods

The table below summarises the principal risks identified in respect of the advertising systems, the mitigations operating against each risk, and the evaluation methods supporting the determination that the mitigations are effective:


[redacted]

[redacted]



Each risk and the corresponding mitigations are addressed in greater detail in the sub-sections below.

[redacted]

[redacted]

Special category data - prohibition on use for ad targeting (Article 26(3) DSA / Article 9 GDPR)

Article 26(3) DSA prohibits providers of online platforms from presenting advertisements to recipients of the service based on profiling using special categories of personal data, as defined in Article 9(1) GDPR. Special categories of personal data include information revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health, or data concerning a natural person's sex life or sexual orientation.

Snap's advertising-system targeting parameters - [redacted] do not include special category data. Snap's Terms applicable to advertisers prohibit advertisers from sending such data from their own properties. Where advertisers are identified as sending special category data to Snap, the data is removed. Targeting on the basis of special categories is not permitted under the platform framework.[138]

[137] Snap Inc., 'Snapchat Ads Transparency', sections 'Information Snap Collects' and 'Controlling the Ads You See' (accessed 4 August 2026); Snap Inc., (internal, current as at 30 June 2026).
[138] Snap Inc., 'General Requirements', section 'Privacy: Data Collection and Usage'; Snap Inc., 'Snap Advertiser Privacy and Security FAQs' (accessed 4 August 2026).

[redacted]

[redacted]

Discrimination - Special Targeting Models

Certain advertisers operate in regulated sectors where rules to prevent discrimination apply - in particular the housing, credit and employment (HCE) sectors.[redacted][39]

[redacted]

Pre-launch testing and mandatory legal and privacy engineering review of any significant changes to these targeting models support the framework on a continuing basis.

Advertising Policies

Snap's Advertising Policies set out the categories of advertisements and targeting practices acceptable on Snapchat. Advertisements must comply with applicable laws and regulations in each geographic area where they will run; the collection of sensitive categories of data is prohibited. The Advertising Policies operate alongside the Community Guidelines to support the integrity of the advertising surface.

As recorded in Section 4 of this Report, advertisers could in principle use the advertising systems to disseminate information that is illegal or that could otherwise harm users, impact fundamental EU rights, or negatively affect public security or health. The Advertising Policies and the advertising-review framework recorded below are the principal operational mechanisms through which this risk is mitigated.

Snap Ads Manager and the systems through which advertisers create and submit advertising have been adapted to require agreement to the Advertising Policies and to provide ready access to guidance on their requirements.

[139] Snap Inc., 'How to Create an Ad Account', section concerning housing, credit and employment advertisements (accessed 4 August 2026); Snap Inc., (internal, current as at 30 June 2026).
[140] Snap Inc., 'Snapchat Ads Transparency', section 'Information Snap Collects' (accessed 4 August 2026).

[redacted]

Advertising Review

Notwithstanding advertiser agreement to the Advertising Policies, advertisers may deliberately or mistakenly seek to publish advertisements that violate the Policies or applicable law. The advertising-review framework operates as the principal pre-publication mitigation against violating advertising.

[redacted]

Reviewer responsibilities

Ad Review team members are responsible for reviewing ad submissions to confirm compliance with Snap's creative policies and technical requirements. The Advertising Policies are applied to assess compliance. Advertisements must comply with both the Community Guidelines and the Advertising Policies in order to be approved. [redacted]

Higher-risk and minor-protection-relevant categories

The review framework includes specific calibration for higher-risk advertisement categories and for minor-protection-relevant content:

[redacted]

[141] Snap Inc., internal advertising-review pipeline data covering 1 July 2024 to 30 June 2026, including final-review and rejection totals and prior-period comparators.

[redacted]

Service-level agreement and monitoring

Snap operates a target service-level agreement of advertising-review completion within 24 hours of advertisement creation by the advertiser. Information on the ad-review process and the SLA framework is publicly available through the Snapchat For Business website. Ad reporting and enforcement data are monitored on a continuing basis to support the determination that the review process is identifying violating advertising at a reasonable and proportionate level.

Reactive advertising reporting

[redacted]

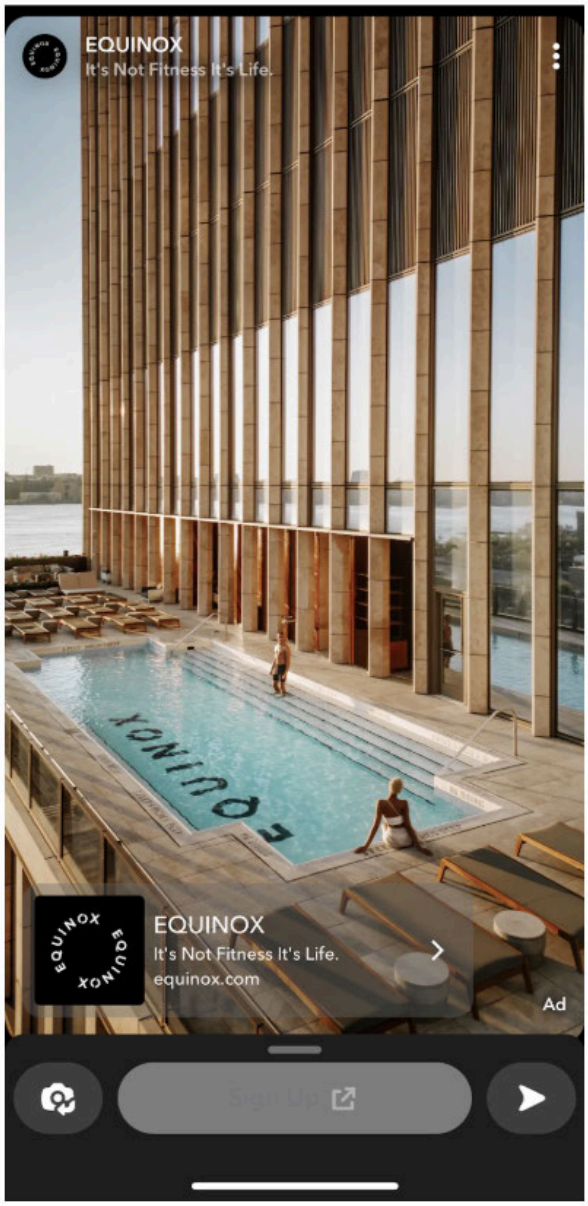
Ad markers and commercial-content transparency (Article 26(1)(a) DSA)

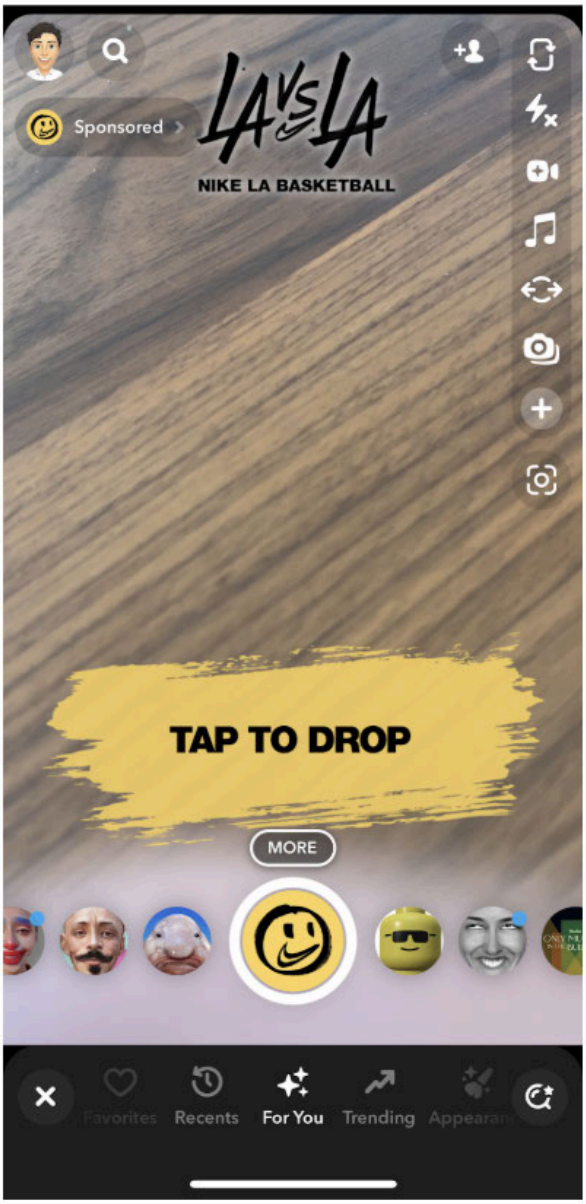
Where users are not aware that content is an advertisement, sponsored or otherwise commercial, there is a risk of confusion, deception or manipulation. The Article 26(1)(a) DSA requirement that recipients of the service be able to identify in a clear, concise and unambiguous manner, in real time, that information is an advertisement is operationalised through the following mechanisms:

- an automatic “Ad” marker is placed on all paid advertisements running on Snapchat;
- Sponsored Lenses are marked “Sponsored”;
- the Commercial Content Policy requires all organic content posted by influencers to be marked appropriately;
- a “Paid Partnership” tag tool is available to influencers and users for posting commercial content, supporting compliance with the Commercial Content Policy and applicable legal obligations under, inter alia, Directive 2010/13/EU (AVMSD) and consumer-protection frameworks.

Ad Marker example

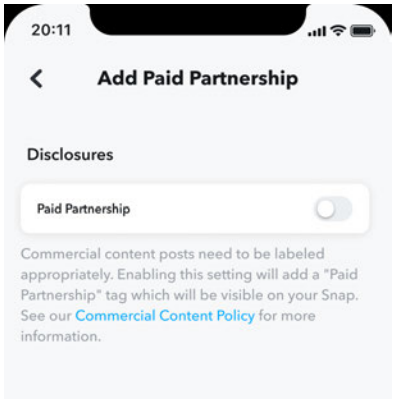
Sponsored Lens Example


EQUINOX
It's Not Fitness It's Life.
EQUINOX
It's Not Fitness It's Life.
equinox.com
Ad
Sign Up

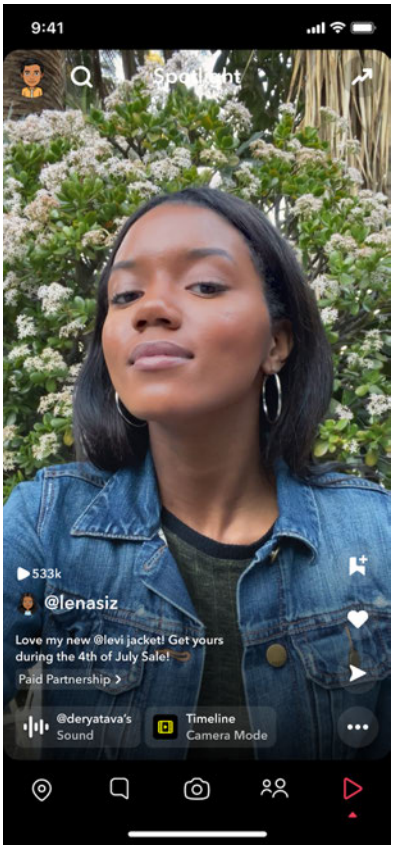

Sponsored
LA VS LA
NIKE LA BASKETBALL
TAP TO DROP
MORE
Favorites
Recents
For You
Trending

[redacted]

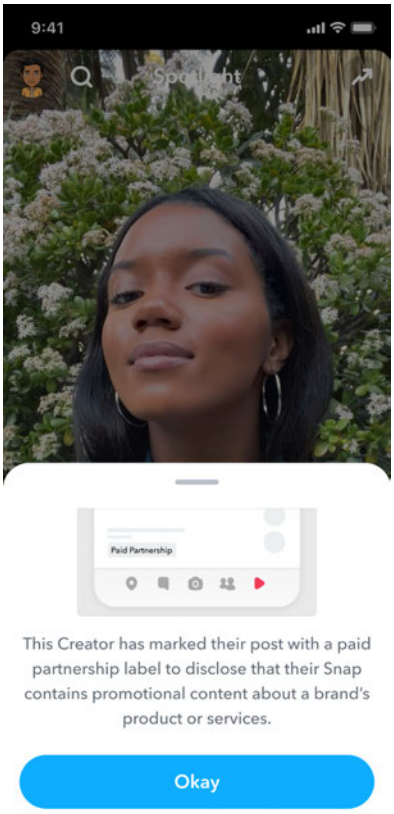
Add Paid Partnership



Paid Partnership Label



Paid Partnership Explainer



Political advertising

Snap discontinued political advertising in the European Union in 2025, in advance of the entry into application of Regulation (EU) 2024/900 on the Transparency and Targeting of Political Advertising in October 2025. The discontinuation removes a documented risk pathway for false or misleading political claims being introduced into the EU advertising surface. The discontinuation is recorded in detail in Section 4.3.1 (Negative Effects on Democratic and Electoral Processes), Section 4.1.10 (Harmful False Information), Section 4.3.2 (Civic Discourse), Section 4.2.7 (Right to Consumer Protection) and Section 4.4.3 (Negative Effects on Minors).

Snap's political advertising library[142], which provides historical disclosure of political advertising delivered through Snapchat, continues to be made available. The library supports both historical accountability for political advertising delivered prior to the discontinuation and ongoing transparency for political advertising delivered in jurisdictions where political advertising remains permitted.

[142] https://www.snap.com/political-ads

[redacted]

Political advertising in approved markets outside the European Union

Outside the European Union, Snap may permit political and advocacy advertising in approved markets, subject to applicable local law and case-by-case review. This includes the United States, except that Snap does not permit advertisements relating to state or municipal elections or ballot initiatives in Washington State. In Canada, Snap does not permit partisan advertising or election advertising as defined by the Canada Elections Act. The political-advertising categories available therefore vary by jurisdiction.[143]

The Advertising Policies contain a sub-section addressing political and advocacy advertising. The framework supports political, election-related, advocacy and issue advertising subject to additional advertiser disclosures and human review:

- political advertising must comply with all applicable laws and regulations, including national election laws, copyright law, and defamation law;
- all political advertising must include a “paid for by” message in the advertisement, followed by the name of the paying person or entity. “Paid for by” disclosure may also be required on advertising content that links to political content, advertising content for political merchandise, or in other cases as appropriate;
- political advertising must comply with the Terms of Service, the Community Guidelines and the Advertising Policies on the same basis as all other advertising;
- political advertisers are encouraged to be positive; however, attack advertising is not categorically banned (campaigning against a candidate or party is generally permissible where consistent with the broader Advertising Policies). Political advertisements must not include attacks relating to a candidate’s personal life;
- political advertising is reviewed on a case-by-case basis. Political advertisers are required to complete the political advertiser form. Snap reserves the right to reject, or to request modifications to, advertisements that violate the standards listed above or that are otherwise inappropriate. Discretion in this respect is not exercised with the intent to favour or disfavour any candidate, political view or political party;
- Snap may publicly display and otherwise disclose information relating to political advertising, including ad content, targeting details, delivery, spend, and other campaign information.

[redacted]

Personal Data Use for Targeting - User Choice

[143] Snap Inc., ’Snap Political and Advocacy Advertising Policies’, sections ’Requirements’, ’Geo-specific Requirements’ and ’Snap’s Rights’ (accessed 4 August 2026). The policy identifies specific restrictions applicable in Canada, Washington State and the European Union, while providing for case-by-case review of political advertising under applicable local law.

[redacted]

Opting out of personalised advertisements

Users in the European Union have the option to disable personalised public content recommendations in Settings under "European Union Controls." Where this option is exercised, Snapchat prevents personalised content from being served to the EU user. Snap's Ad Server automatically prevents targeted ad types (including pre-defined audiences and custom audiences) from being approved for ads reaching Snapchatters who have opted out. Where personalisation is disabled, public content is recommended on the basis of basic data only (the language set on the device, the user's age, and the user's country); content remains available but with reduced personalisation.

Do Not Track

Snapchat respects the device iOS App Tracking Transparency setting regardless of user Ad Preference choices.

Ad Preferences

EU users who have enabled personalised content have additional Ad Preferences available for modification within Snapchat. Users have choices over how Snap and its third-party advertising partners use information about them for advertising purposes both on and off Snapchat. The settings allow users to select whether to receive Audience-Based ads, Activity-Based ads, Ads from Third-Party Ad Networks, or none of these ad options. The settings on this page are automatically disabled for users under 18 in the EEA or UK.

[redacted]

Transparency and Control

[redacted]

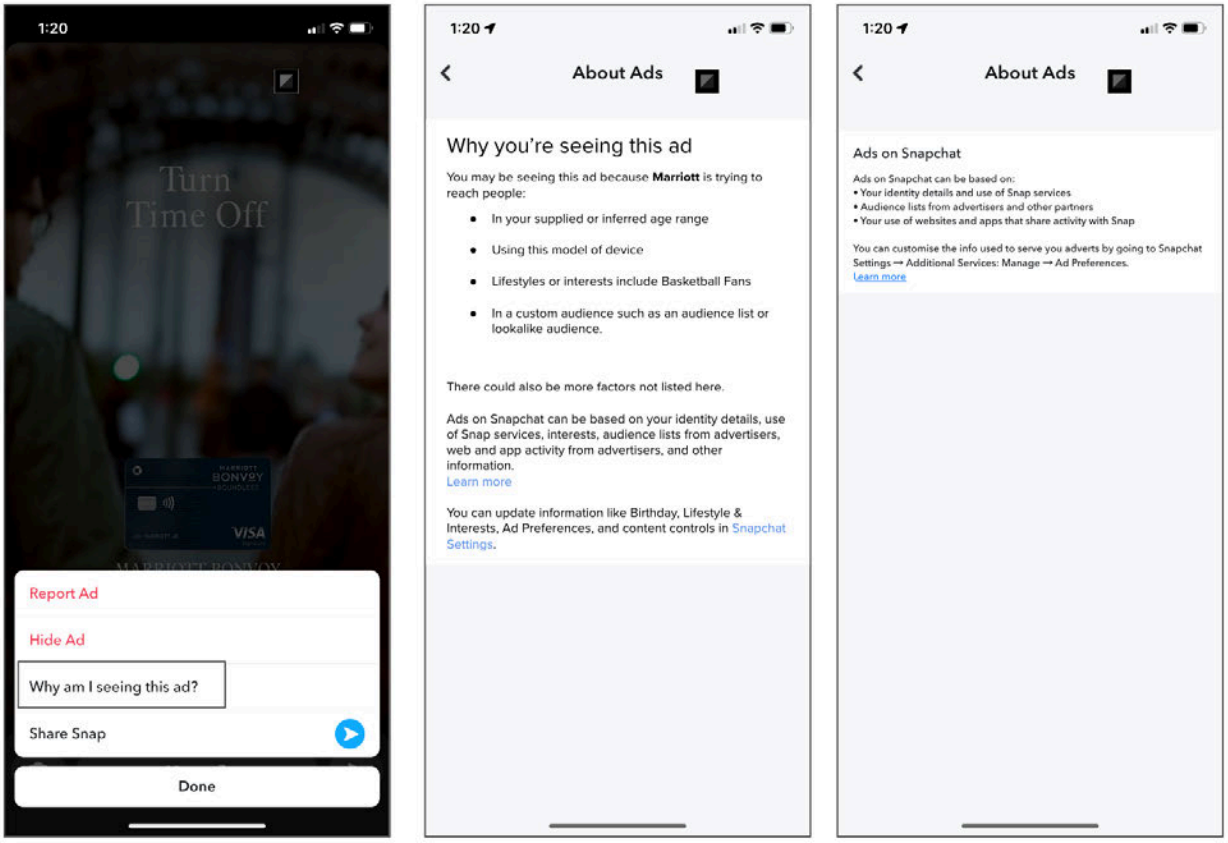
As recorded in Section 5.3 (Transparency), Snap's Privacy Center provides extensive information regarding the processing of personal information, including a dedicated page explaining how personal data is used for advertising. [redacted]

[redacted]

[redacted] All users can restrict the use of third-party data and inclusion in advertiser-supplied audience matches for ad targeting. Pre-launch testing and the ad-review process support the operational performance of these controls.

"Why am I seeing this ad" - Article 26(1)(d) DSA

Article 26(1)(d) DSA requires that recipients of the service be able to access meaningful information directly and easily from the advertisement about the main parameters used to determine the recipient to whom the advertisement is presented and, where applicable, about how to change those parameters. Snapchat operationalises this requirement through the "Why am I seeing this ad" functionality, which is available to Snapchatters in the EU. Selecting "Why am I seeing this ad" presents additional information on the targeting parameters for the advertisement in question.[145]



Ads Library / Ads Gallery (Article 39 DSA)

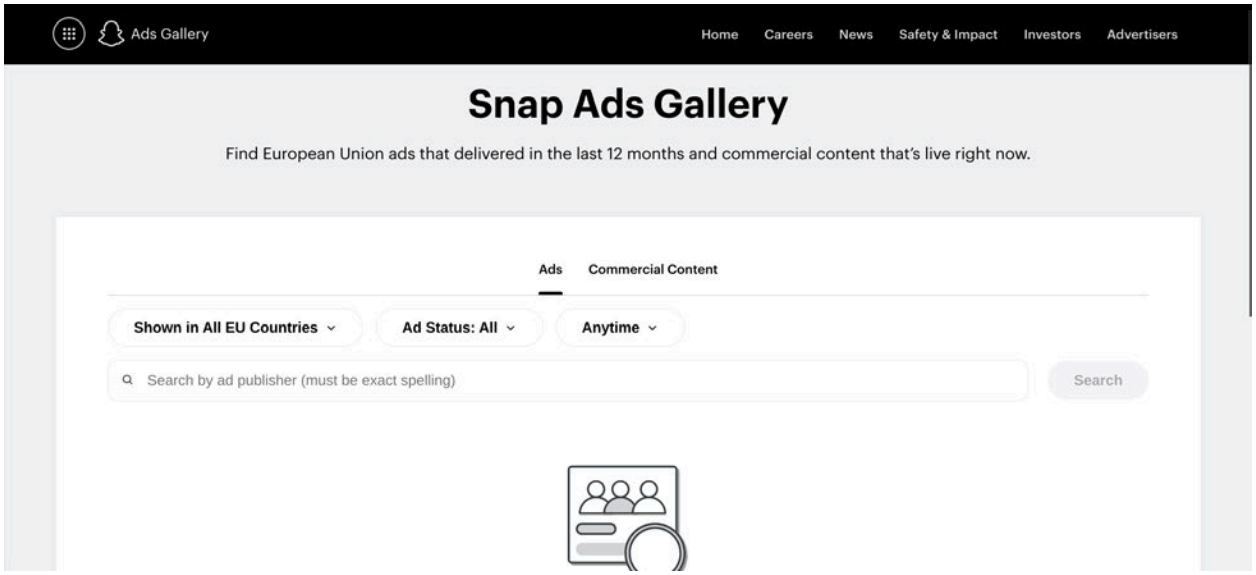
[144] Snap Inc., 'Snapchat Ads Transparency', section 'Information Snap Collects'; Snap Inc., 'How Location May Affect Ad Preferences Settings'; Snap Inc., 'Frequently Asked Questions About the European Digital Services Act' (accessed 4 August 2026).

[145] Snap Inc., 'Frequently Asked Questions About the European Digital Services Act' (accessed 4 August 2026); Snap Inc., 'Updating our Advertising in Europe and the UK' (24 July 2023).

Article 39 DSA requires VLOPs that present advertisements on their online interfaces to compile and make publicly available, through a searchable and reliable tool with multi-criteria queries and through application programming interfaces, a repository of advertising data covering the entire period during which an advertisement is presented and for one year after the advertisement was last presented.

Snap Ads Gallery

The Snap Ads Gallery operates as the principal compliance mechanism for Article 39 DSA. The Gallery makes accessible the historical and current advertisements that have been delivered to Snapchat users in the EU and in other jurisdictions where the framework operates. The Gallery is publicly available, searchable and filterable by a range of pre-defined parameters (including country targeted and advertiser name), and includes an API interface for programmatic access. The Gallery maintains advertisements for one year after the advertisement last airs. The Commercial Content section of the Gallery includes links to all live organic content marked with the "Paid Partnership" tag (Article 39(2)(f) DSA).[146]



Article 39 DSA disclosure mapping

The mapping below identifies the operational implementation of each Article 39(2) DSA disclosure requirement in the Snap Ads Gallery:

Article 39 DSA disclosure requirement	Operational implementation in the Snap Ads Gallery

[146] Snap Inc., 'Ads Gallery API', 'Introduction' and 'Using the API' (accessed 4 August 2026); Snap Inc., 'Frequently Asked Questions About the European Digital Services Act' (accessed 4 August 2026).

Content of the advertisement (Art. 39(2)(a))	"Ad Creative" field in the Main Ad Modal: ad content, including the name of the product, service or brand and the subject matter of the advertisement.
Person on whose behalf the advertisement is presented (Art. 39(2)(b))	"Brand Advertised" field in the Main Ad Modal: the natural or legal person on whose behalf the advertisement is presented.
Person who paid for the advertisement (Art. 39(2)(c))	"Ad Publisher" field in the Main Ad Modal: the natural or legal person who paid for the advertisement.
Period during which the advertisement was presented (Art. 39(2)(d))	"Ad Start Date" and "Ad End Date" fields in the Main Ad Modal.
Whether targeted to particular groups and the main targeting parameters (Art. 39(2)(e))	"Targeted devices and demographics" field in the Ad Details Modal: whether the advertisement was intended to be presented specifically to one or more particular groups of recipients (devices and demographics). Exclusion targeting is not supported.
Article 26(2) DSA-identified commercial communications (Art. 39(2)(f))	Commercial Content section of the Ads Gallery includes links to live organic content marked with the "Paid Partnership" tag.
Total number of recipients reached, with Member State breakdown (Art. 39(2)(g))	"Total Impressions" field in the Main Ad Modal; "Impressions by Member State" field in the Ad Details Modal (where zero recipients were reached, the ad does not appear).

Advertisements that were delivered and subsequently taken down are marked as Rejected in the Ads Gallery, supporting transparency on enforcement outcomes. The Snap Ads Gallery is maintained by the Ads API and Ads Formats teams. The library underwent pre-launch testing and continues to be developed on the basis of further testing.

Freedom of Expression dimension

The purpose of Snap Ads is to amplify advertisers' commercial messages. The content of advertisements is, in the ordinary course, commercial rather than political or expressive. Article 11 EU Charter freedom-of-expression considerations therefore engage with reduced intensity in the advertising context: while a right to speak applies to commercial speech, the right does not extend to algorithmic amplification on the platform. Substantive cross-references on the freedom-of-expression treatment of advertising are recorded in Section 4.2.2 (Right to Freedom of Expression).

[redacted]

Material change for the 2026 assessment period

The principal change to the advertising-systems framework relative to the 2025 Report is the discontinuation of EU political advertising in 2025, ahead of the entry into application of Regulation (EU) 2024/900 in October 2025. The discontinuation:

- removes a documented risk pathway for false or misleading political claims being introduced into the EU advertising surface;
- aligns Snap's framework with the EU Political Advertising Regulation;
- preserves continuity of the political advertising library for historical disclosure of advertising delivered prior to the discontinuation, and for advertising in jurisdictions where political advertising remains permitted.

Operationally, [redacted] distinct advertisements completed final review during the 2026 assessment period, an increase of approximately [redacted] from the prior period, while [redacted] advertisements were rejected, a decrease of approximately [redacted] Further detail is recorded in Section 4.1.11. The substantive Article 26 DSA advertisement-transparency framework, the Article 39 DSA Ads Gallery, the audience-minimum requirements, the special-targeting controls, the prohibition on advertising-targeting based on minor users' activity data, and the personalised-ads-off-by-default position for users under 18 in the EEA and UK are unchanged in their core architecture.

Conclusion

Targeted advertising on Snapchat supports the maintenance and development of the service for users regardless of their ability to pay. The framework recorded in this Section operationalises the targeted measures adopted under Article 35(1)(e) DSA against the systemic risks identified across Section 4. The integrated set of design, policy, review, transparency and user-control measures supports the determination that the framework operates in a manner that is reasonable, proportionate and effective.

The principal architectural commitments of the framework - [redacted] the Article 26(3) DSA / Article 9 GDPR special-category-data prohibition; the HCE special-targeting controls; the prohibition on advertising-targeting based on minor users' activity data; the personalised-ads-off-by-default position for users under 18 in the EEA and UK; the EU political advertising discontinuation in October 2025; the Article 26 DSA advertisement-transparency framework operationalised through the "Ad" markers, the "Why am I seeing this ad" functionality, and the Article 39 DSA Ads Gallery - together provide an integrated regulatory-aligned response to the advertising-system risk profile.

As recorded in Section 4 of this Report, the framework recorded in this Section, in combination with the other mitigations recorded across Section 5, supports the conclusion that Snap's

[redacted]

mitigations for the systemic risks identified in respect of the in-scope services are reasonable, proportionate and effective. There is no change in this overall conclusion from the 2025 Report. The framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes during the 2026 assessment period and beyond.

5.8 Protection of Minors

Introduction

Section 5.8 sets out the integrated minor-protection mitigation framework operating across Snapchat's in-scope services. The framework is the principal substantive place for the minor-protection commitments mapped to Article 35(1)(j) DSA across Section 4 of this Report, and is the operational home for Snap's response to Article 28 DSA, the European Commission's Article 28 Guidelines published in 2025.

The framework operates through a multi-layer architecture spanning app-store, device, platform and product-level safeguards, governed by the six tenets recorded below and operationalised through the Age Appropriate Design Code (AADC) framework. The architecture reflects the foundational positioning of privacy, safety, and security at the core of Snap's product design and the long-standing approach of designing for the youngest users first - with the protections so designed forming the baseline safety layer applied across the user community.

Snapchat had an average of [redacted] monthly active recipients aged 13–17 in the European Union during the 2026 assessment period, representing approximately [redacted] of its total EU monthly active recipient population. The minor-user population is therefore both substantial in absolute terms and a discrete protected population to which the elevated mitigation framework recorded in this Section applies.

Administration and Oversight

[redacted]

[redacted]

The Six Tenets Framework

From the outset, Snapchat was designed to take into consideration the rights recognised under the UNCRC and to operate in the best interests of the child (with foundational design choices including data deletion by default and safety and privacy by design and default as the baseline). The framework operates through six tenets that guide product, policy, and operational decision-making in respect of minor users:

Tenet	Operational expression
1. Act in the best interests of teens	[redacted]

[redacted]

2. Strict default settings for all users	[redacted]
3. Respect teens' freedom to express themselves safely	[redacted]
4. Recognise teens' right to information about the world	[redacted]
5. Position parents/guardians to help guide teens	[redacted]
6. Attach a heightened safety interest to teens	[redacted]

[redacted]

[redacted]

[redacted]. Privacy and Safety by Design principles support risk mitigation for minors. [redacted] Snapchat's features are subject to AADC assessments considering the fifteen standards of the Code, conducted with product teams across the business.

Age Appropriate Design Code (AADC)

Snap has adopted the Age Appropriate Design Code (AADC) guidance for developers to incorporate appropriate and proportionate measures to safeguard the privacy, safety and security of users aged 13-17 within platform products and features. Snap aligns its practices with principles reflected in age-appropriate design frameworks, including the United Kingdom's Children's Code and relevant European and other comparable frameworks. The AADC framework consists of fifteen core standards, with Snap's operational expression recorded below:

AADC core standard	[redacted]
1. Best interests of the child	
2. Data protection impact assessments	
3. Age-appropriate application	
4. Transparency	
5. Detrimental use of data	

[redacted]

6. Policies and community standards	[redacted]
7. Default settings	
8. Data minimisation	
9. Data sharing	
10. Geolocation	
11. Parental controls	
12. Profiling	
13. Nudge techniques	
14. Connected toys and devices	

[redacted]

15. Online tools	[redacted]

[redacted] The Commission has published its guidance on Article 28; [redacted]. The Article 28 Guidelines mapping is recorded below.

Article 28 DSA Guidelines (Commission Guidance, 2025)

Following the publication by the European Commission of its Guidance on Article 28 DSA in 2025, [redacted]. The detailed assessment is recorded below, mapped against the seven principal sections of the Guidance.

General Principles	[redacted]
Risk Review	

[redacted]

Service Design - Age Assurance	[redacted]
Service Design - Defaults and Account Settings	
Service Design - Interfaces, Tools and AI	

[redacted]

Service Design - Recommender Systems and Content Moderation	[redacted]
Service Design - Commercial Practices	

[redacted]

Reporting Support and Guardian Tools	[redacted]
Governance	
Transparency to Minors	

[redacted] Effective industry-wide standards for assessing a high level of privacy, safety and security for teens, in line with existing AADC frameworks, DPIA requirements and Privacy and Safety by Design obligations, are an objective Snap supports. Further guidance and regulation should consider, in particular, the role of online platforms that are gateways (such as device operating systems, app stores and web browsers) through which parents and teens engage with such platforms. US

[redacted]

states (including Utah and Texas) have enacted legislation requiring app stores to conduct age verification and to share verified age signals with app developers. [redacted]

Age Assurance Framework

Snap's age-assurance framework operates through a series of complementary controls across the user lifecycle:

1. [redacted]

Multi-layer architecture

Snap's age-assurance framework operates through a multi-layer architecture, with each layer providing complementary signals and protections:

Layer	Description

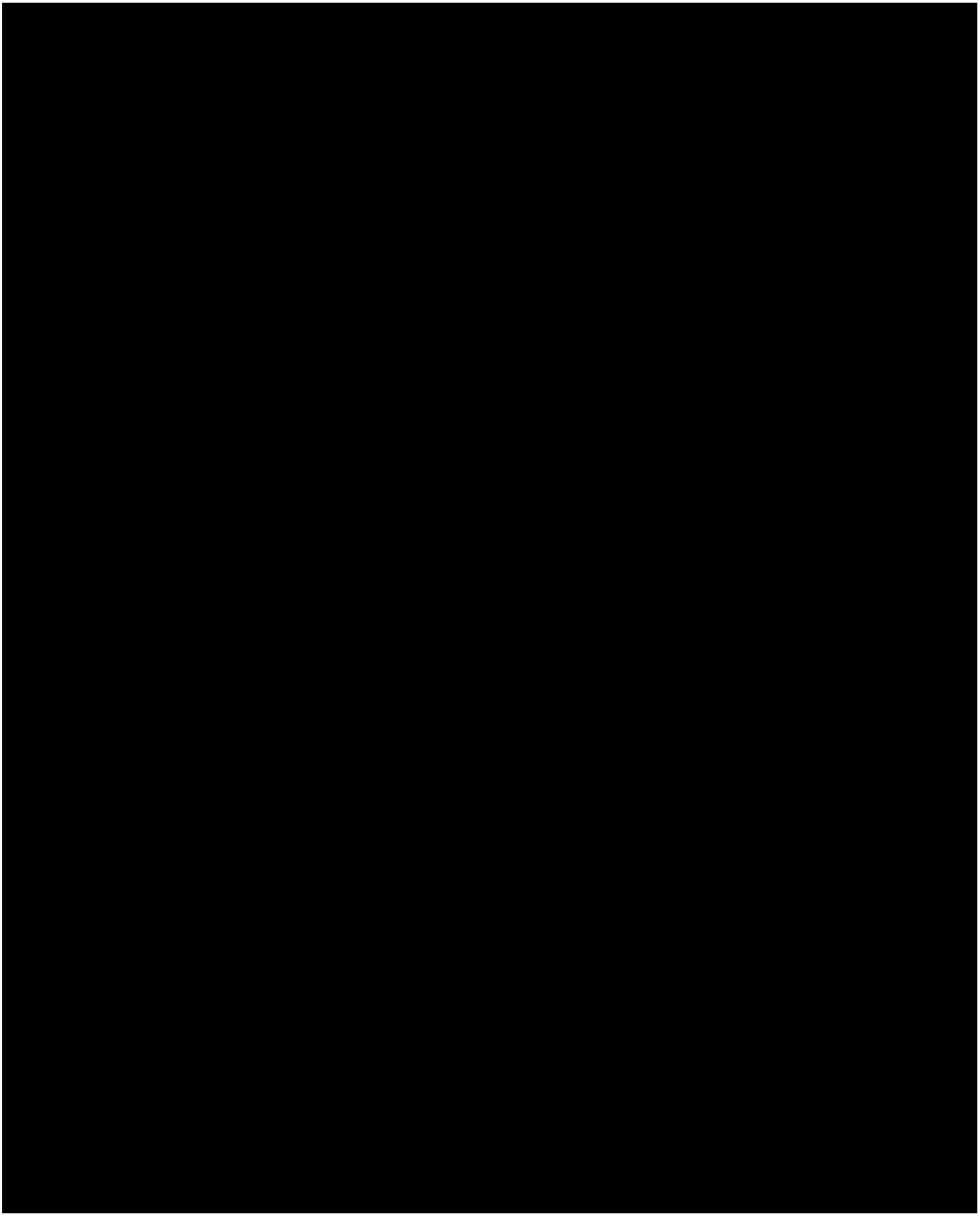
[redacted]

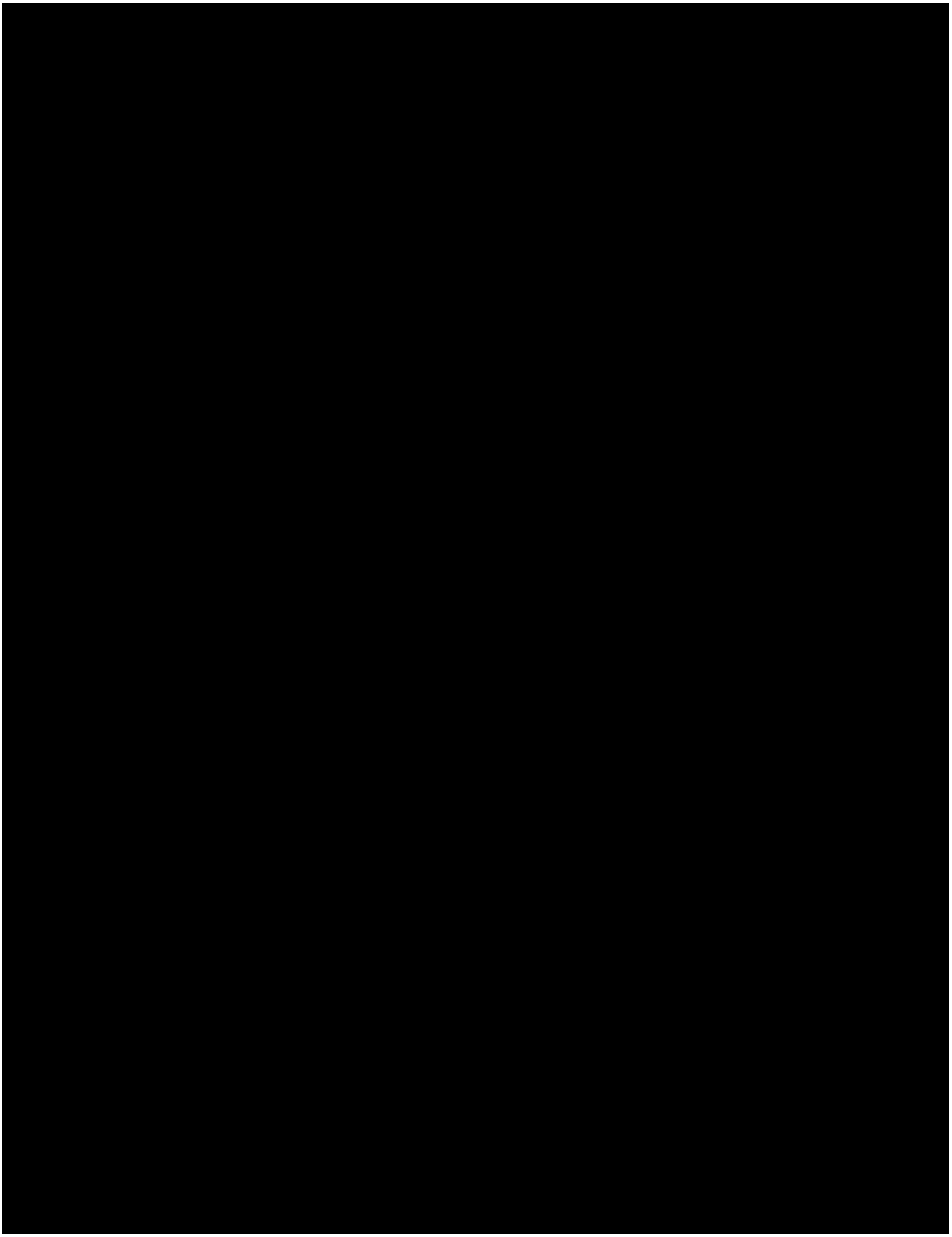
[redacted]

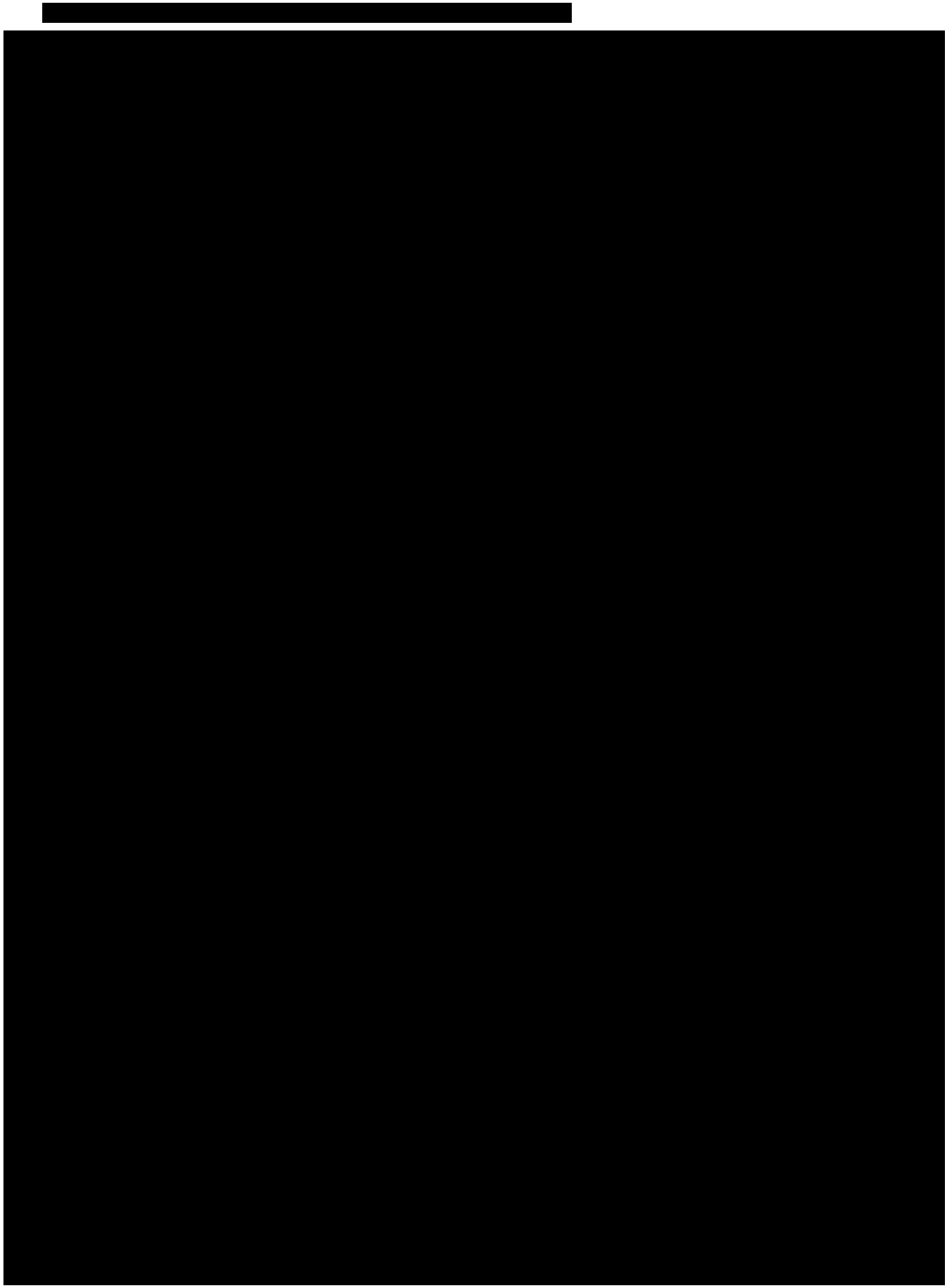
[redacted]

Registration and access

[redacted]







[redacted]

[redacted]

Child Safety Working Group oversight

The Child Safety Working Group is a dedicated cross-functional group overseeing Snap's age-assurance efforts. The group consists of more than ten full-time employees with representatives from product, privacy legal, product legal, policy, and trust and safety teams. Iterations of the group activate as needed to meet in working groups focused on exploring, discussing and assessing concepts and solutions related to age assurance through risk assessments, discussions with the product team, and collaboration with external legislators, regulators, peer platforms, vendors, experts, NGOs and other stakeholders.

[redacted]

Ongoing evaluation

Snap continues to evaluate the framework and to consider possible concepts and approaches with industry peers and third-party age-assurance vendors, supporting alignment with industry practice. Snap has supported legislators and NGOs in the United Kingdom, France and elsewhere in the EU on enhancing the role of app stores, online devices and web browsers in providing appropriate interfaces for age assurance and parental controls. The framework continues to focus on the trade-offs that age-assurance measures present: the trade-off between safety of minors and compromise of user privacy/data security; the accuracy and reliability of age-estimation technology (particularly for younger or ethnic-minority populations); the fairness of methods that may disadvantage users without official government IDs or bank accounts; and the harm to industry competitiveness from the cost of adopting third-party technology at scale.

Children's Advertising Review Unit (CARU) recognition has been received by Snap for measures protecting under-age users. Reassessment continues on a continuing basis, including

[151] Snap Inc., Response of 25 June 2026 to the European Commission's Request for Information, responses to Questions 7, 12 and 15.

[redacted]

engagement with industry partners, regulators, and third-party age-assurance vendors. [redacted]

Multi-Layer Safeguards Architecture

Beyond the foundational defaults that apply to all users, additional safeguards apply at the app-store, device, platform and product levels. The product-level safeguards are recorded in detail below.

App Store Level

App stores host a vast range of apps presenting a wide array of risks and represent a higher-risk situation in the digital ecosystem. App stores can generate higher and exponential risks for a minor than the risks created by individual services. Larger app stores recognise this risk and apply age-gates that prevent users from downloading apps if their app-store account age is below the app-provider-specified minimum (and where applicable parents via the device operating system's account-level family controls). App stores typically rely on the app-store account's declared age (often the same account used by the device operating system).

Improvements that could be made by app-store providers for the benefit of the entire ecosystem include: (i) stronger age assurance at the app-store / device-OS account level and sharing that signal with developers to support minor-protection measures; and (ii) allowing developers to set a more precise minimum age (the classical 12+ category does not align with the 13+ threshold commonly specified by application services).

Device Level

Many devices now come with well-being settings, including bedtime modes that turn off device and app notifications and turn the screen black-and-white to encourage sleep. Operating-system family controls (Apple Family Sharing, Google Family Link, Microsoft Family Safety) allow parents and other responsible adults to view usage, set time limits, and disable per-app access. The device-OS family-control providers are also gatekeepers under the EU Digital Markets Act, providing deeper levels of visibility and control for their own first-party services. This level of interoperability and access would be useful for Snap's own Family Center, which could increase awareness and accessibility for parents and other responsible adults who do not have a Snapchat account themselves. Snap continues to advocate for further multi-stakeholder dialogue to drive solutions providing equal access and interoperability across the industry.

Platform Level

[redacted]

[redacted]

[redacted]

Product Level

Surface-specific safeguards operate across each in-scope product. Detail by surface is recorded below.

Lenses

Snap takes steps to help teens understand the design and functioning of Lenses and the risks inherent in Lenses that may cover controversial topics (for example, appearance-altering Lenses). Snap rejects Lenses that intentionally imitate specific cosmetic procedures and applies strict age and distribution limits to advertising or sponsored Lenses promoting cosmetic procedures. Snap allows cartoon-like Lenses (with whimsical, non-photo-realistic features) on the platform.

Context cards attach to sending or posting Snaps so the receiver knows that a Lens has been applied. The Baby Lens is an example of a Lens that can transform an adult user's appearance to look like a baby; concerns about adults using such a Lens to pose as younger users to groom teens were addressed by adding a rattle to the Lens creative to clearly indicate it is a Lens transformation, and by ensuring the context card indicates to on-platform viewers that a Lens has been applied to the Snap.

Snap geofences or age-gates certain Lenses (for example, related to alcohol, gambling, NFTs, and other regulated goods). For example, a Vivino Lens is shown only to 18+ users in the EU, mitigating exposure to such content for users under the legal drinking age. Snap respects teens' right to privacy in the context of Lenses: Lenses rely on object detection rather than facial identification (Lenses can determine whether a face is present, but do not identify specific faces), and data collected from Lenses is used only to provide the service in line with Snap's privacy commitments.

[redacted]

Public Profiles

[redacted]

[redacted]

Public Content - Creating

[redacted]

Public Content - Viewing

[redacted]

Discover

Discover policies are written with the understanding that people as young as 13 may be viewing the content, and that content should be age-appropriate. Where content is not appropriate for 13+, partners have tools to mark their content 18+ at their discretion. Regulated goods, R-rated trailers, ESRB 17+-rated games, cosmetic surgery and supplements-based weight-loss products, dating apps, and provocative or sexualised imagery do not appear as ads to teens. Discover Publisher partners must adhere to the Commercial Content Policy, which requires content to be appropriate for a 13+ audience or to be age-gated at 18+ where appropriate.

Spotlight

Spotlight content is moderated to prevent sexually-suggestive content from appearing to teens. [redacted] Content scored above the suggestive threshold is removed from teen-recommended feeds.

Specific Spotlight protections for teens include:

[redacted]

[redacted]

[redacted]

Snap Map

Snap Map opens in Ghost Mode (location-sharing off) by default. [redacted]

Advertising

Detail on the advertising-side minor-protection framework is recorded in Section 5.7 (Advertising Systems). [redacted] the discontinuation of EU political advertising as of October 2025.

Reporting and Blocking

In-app reporting

[redacted]

Blocking and removing friends

All users have the option to block another user. Blocking a user prevents the friend from viewing friend content posted by, or sending Snaps and Chats to, the blocking Snapchatter. [redacted]

[redacted]

[redacted]Removing a friend from the friends list prevents the removed Snapchatter from viewing friend content and, by default, from chatting or Snapping with the user.

Private Messaging Default messaging restrictions

[redacted]

In-Chat warnings

[redacted] Substantive detail on the friending-protection refinements is recorded in Section 5.4.4.

Media scanning

[redacted]. Substantive detail is recorded in Section 4.1.1.

Friending

Search and Find Friends visibility

Teens appear in another user's Search results only in limited circumstances - [redacted] Similar protections apply to prevent teens from searching for unknown adults. [redacted]

[redacted]

Default friending architecture

[redacted]

[redacted]

[redacted]

Friend Check-Up and in-app warnings

Friend Check-Up prompts Snapchatters to review their friend lists and remove friends with whom they are no longer in contact, supporting an up-to-date network focused on close friends. Pop-up warnings are provided: [redacted]

Family Center / Parent Tools

Family Center is Snap's in-app parental supervision tool. It provides parents, caregivers and other trusted adults with visibility into the teen's friends list and recent communications (last seven days), and with the ability to:

- view the teen's friends list;
- see which friends the teen has recently communicated with (last seven days);
- restrict the teen's access to Spotlight and Discover content tagged as sensitive by the moderation team;
- disable the teen's ability to engage with the My AI chatbot;
- see who the teen is sharing location with, and request location sharing from the teen (which the teen must approve);
- report concerns about accounts that may violate the Community Guidelines;
- control whether the teen has access to sensitive content on For You (Discover) and Spotlight feeds;
- access helpful resources directly in the app.

Design and balance

Family Center was designed with the objective of supporting both caregivers and teens, balancing parents' interest in greater insight with teens' interest in autonomy and privacy. Notably, teens' messages remain private and are not visible to parents. Consistent with the Guidance's position that compliance with Article 28(1) DSA must never rely exclusively on tools for guardians, and that such tools should be treated as complementary to (not a substitute for) safety-by-design and default measures, Family Center operates as one layer within Snap's broader multi-layer minor-protection architecture rather than as a standalone safeguard. Snap recognises that the effectiveness of guardian tools may be limited where the identity or legal authority of the participating adult is not verified. [redacted] It operates as a complementary layer within Snap's broader

[redacted]

safety-by-design framework. [redacted]

Ongoing improvements to Family Center accessibility

During the 2026 assessment period, Snap continued to evaluate improvements to the accessibility and discoverability of Family Center, [redacted]

Snap will continue to assess opportunities to increase awareness and uptake of Family Center while preserving teen-side consent and the complementary role of guardian tools within the broader safety-by-design framework.

External recognition

In October 2022, Jugendschutz.net, the joint competence center of the German Federal Government and the German states for the protection of children and young people online, reviewed the launch of Family Center. It noted that Family Center may support conversations between caregivers and teens about online contacts and positively highlighted the transparency provided to teens, including the requirement for teen-side consent before supervision is activated.[152]

Three-Pillar Operational Framework

Snap's minor-protection mitigation framework operates around three core pillars (recorded also in Section 4.4.3, Negative Effects on Minors):



[152] Jugendschutz.net, 'Snapchat: "Family Center" in der App verfügbar' (19 October 2022).

[redacted]

Transparency to Minors

Information regarding Snapchat services is provided in a manner that is clear and comprehensible across age groups. Readability of key Terms and Conditions and Privacy Notices is verified using automated readability tools. The vast majority of these documents are confirmed to be understandable for users. The main Terms and Conditions document is, necessarily, a formal legal document with longer provisions and more complex language; to support readability for younger users, the Terms have been organised into sections with appropriate flow, succinct section headers, and short summaries at the bottom of each section. The short summaries have been tested for readability and confirmed to be understandable for users including teens. Short explainers have been provided for Community Guidelines to support user understanding.

The Privacy Center was designed for the youngest Snapchatters and was intentionally developed to be easy to read and understand by all members of the community. The Privacy and Safety Hub (with pages such as Privacy by Product) provides high-level summaries of privacy and safety practices across products and features. A video resource visualises the privacy practices, and icons and other best practices recommended by privacy and safety experts and the AADC frameworks are used throughout.

The substantive minor-protection mitigation framework (the six tenets; the AADC framework; the multi-layer architecture; the inferred-age model; the proactive public Profile language-scanning and LLM-assisted review; the Family Center; the product-level safeguards across Lenses, Public Profiles, Spotlight, Discover, Snap Map and Advertising) is unchanged in its core architecture.

Conclusion

Snap's commitment to the protection of teens on Snapchat is operationalised through the integrated multi-layer minor-protection framework recorded in this Section. The framework is built on the six-tenet foundation: act in the best interests of teens; strict default settings for all users; respect teens' freedom to express themselves safely; recognise teens' right to information about the world; position parents and guardians to help guide teens; attach a heightened safety interest to teens. These tenets are operationalised through Snap's alignment with principles reflected in age-appropriate design frameworks, including the United Kingdom's Children's Code and relevant European and other comparable frameworks; the multi-layer architecture (app-store, device, platform and product); the Family Center oversight framework; the integrated PASS Review and DDIA framework; engagement with the Safety Advisory Board, Council for Digital Well-Being and Teen Council; and continuing engagement with Member State child-protection authorities and civil-society organisations.

As recorded in Section 4 of this Report, and consistent with the determinations recorded across Sections 4.2.6 (Children's Rights), 4.4.3 (Negative Effects on Minors), and 4.4.4 (Physical and Mental Well-Being), the targeted measures recorded in this Section to protect the rights of the

child, including age-assurance and parental-control tools, in combination with the other mitigations recorded in Section 5, are reasonable, proportionate and effective mitigation measures for the risks identified for Snapchat's in-scope services. Snap regards the 2026 changes as the natural next iteration of its continuous review (rather than evidence that the prior risk determination required revision). The overall risk classification remains stable, even as the specific mitigations supporting it continue to evolve and be refined in light of operational data and external guidance.

Snap has actively supported the European Commission and others in establishing guidance on a high level of privacy, safety and security for the purposes of Article 28 DSA. The framework continues to engage with the Commission's published Article 28 Guidelines, and will continue to be subject to ongoing monitoring and refinement through the Section 6 risk-detection processes during the 2026 assessment period and beyond.

5.9 Content Authenticity

Introduction

Section 5.9 sets out the content-authenticity framework operating in respect of Snapchat's in-scope services, addressing the Article 35(1)(k) DSA mitigation category. There continues to be intense interest and regulatory attention surrounding the ways in which advancements in generative AI technologies are affecting online platforms. Content authenticity is a challenging topic without a single technical solution, and the framework recorded below is structured as a multi-layered approach rather than as reliance on any single mitigation pathway.

Snap recognises that AI-generated or AI-transformed content may be distributed through Snapchat's in-scope services (notably Spotlight and Discover), as the Community Guidelines explicitly acknowledge. Such content, like any other content distributed on the in-scope services, may constitute illegal content or otherwise violate the Snap Terms, and may engage the systemic risks identified under Article 34 DSA. Snap continues to monitor developments and industry practice on a continuing basis, including in respect of the appropriate use of prominent markings and other measures to distinguish content that falsely appears to be authentic.

The 2026 assessment period sees substantial evolution in the EU regulatory framework bearing on this Section, in particular through the staged entry into application of Regulation (EU) 2024/1689 (the EU AI Act), the operation of the General-Purpose AI Code of Practice, and the transition of the previously-voluntary Code of Practice on Disinformation to operate under the

DSA framework as of 1 July 2025. The framework recorded below reflects this evolving regulatory environment.[153]

Risk Assessment Results

Generative AI-related risks are addressed across multiple Category 1, Category 3 and Category 4 risks assessments of Section 4. The principal cross-references and 2026 determinations are summarised below:

Underlying risk	2026 determination and content-authenticity intersection
Section 4.1.10 - Harmful False Information	Lowest likelihood; significant harm industry-wide; Level 3. The substantive treatment recognises that disinformation, deceptive content and "deep fakes" have received significant attention in media, academic and political debate. The likelihood determination on Snapchat's in-scope services is based on Policy Violating Prevalence figures [redacted] at the displayed precision) and the very low VVR figures across surfaces and audiences.
Section 4.3.1 - Democratic and Electoral Processes	Lowest likelihood; severe harm industry-wide; Level 3. The substantive treatment recognises that online platforms may negatively affect electoral processes through the amplification of digital disinformation or deceptive content (whether AI-generated or otherwise). Independent reports of electoral interference on Snapchat have been very limited; engagement during 2024 with the Election Integrity Partnership identified one single incident report from external researchers monitoring risks on Snapchat.
Section 4.3.2 - Civic Discourse	Lowest likelihood; severe harm industry-wide; Level 3. The substantive treatment recognises that platforms may contribute to civic-discourse harm through echo-chamber dynamics, amplification of dis/misinformation, and amplification of extreme or sensational content. The empirical findings on Snapchat's in-scope services do not identify civic-discourse harm as elevated risk.
Section 4.1.1 - CSAM	Lowest likelihood; severe harm; Level 1 (Section 3.5 deviation). Generative AI as a vector for CSAM is recognised as an emerging concern; the substantive treatment in Section 4.1.1 records the operational measures in place against synthetic CSAM (whether authentic or AI-generated).

[153] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence, Articles 50 and 113; European Commission, 'General-Purpose AI Code of Practice now available' (10 July 2025); European Commission, 'The Code of Conduct on Disinformation' (13 February 2025).

[redacted]

Section 4.4.2 - Gender-Based Violence	Lowest likelihood; serious harm; Level 2 (Section 3.5 deviation up). Technology-facilitated GBV (including deepfake NCII and AI-manipulated harassment imagery) is recognised as an emerging risk dimension and is addressed through the integrated Category 1 mitigation framework alongside the content-authenticity measures recorded here.
Section 4.4.3 - Negative Effects on Minors	Lowest likelihood; severe harm; Level 1. Generative AI risks bearing on minors include synthetic minor-targeted content, deepfake material affecting minors, and the use of generative AI by minor users (with the additional safeguards recorded in Section 5.8).

The empirical findings recorded across the underlying risks support the determination that, while content authenticity is a significant area of regulatory and societal attention, the prevalence of substantively harmful AI-generated content on Snapchat's in-scope services has remained low across multiple reporting cycles. The framework recorded below operates to maintain this prevalence trajectory and to support continued resilience as the underlying technologies evolve.

Mitigation Architecture

[redacted].

Snap maintains substantive policies applicable to both the dissemination and the creation of generative AI content. Notwithstanding that the prevalence of harmful false information, fraud and spam, and impersonation on Snapchat's in-scope services remains low (Sections 4.1.10 and 4.1.11), an integrated mitigation framework is operated to support continued low prevalence and to address the evolution of the underlying technologies.

Creation track - Internal Generative AI Policies

Snap operates internal policies governing the development of generative-AI features (such as MyAI).[redacted]

- [redacted]

[redacted]

- [redacted]

Note that on-platform features for creating generative AI content are not, in the ordinary course, part of Snap's in-scope services for the purposes of this risk assessment (with the exception of certain commonplace ad-creation tools). Risks and mitigations relating to these features are assessed through the privacy and safety by design product reviews recorded in Section 6.3 [redacted] The discussion in this Section accordingly addresses the policies as they operate in respect of the matters bearing on the in-scope services.

[redacted]

[redacted]

[redacted]

[redacted]

Dissemination track - Community Guidelines and Terms

In the context of dissemination of content on Snapchat's in-scope services, the Community Guidelines and the Terms of Service set out the rules on what content is allowed on Snapchat. The rules are content-format and tooling-agnostic: they apply to all content formats across the platform, including AI-generated content, and address content irrespective of whether it was generated using generative AI, conventional editing tools (such as Photoshop), or other means.

The Community Guidelines specifically note: "We implement safeguards designed to help keep generative AI content in line with our Community Guidelines, and we expect Snapchatters to use AI responsibly. We reserve the right to take appropriate enforcement action against accounts that use AI to violate our Community Guidelines, up to and including the possible termination of an account."[154]

Detailed rules on civic-discourse and electoral-process content engage the dissemination track. The Community Guidelines prohibit spreading false information that causes harm or is malicious - such as denying the existence of tragic events, unsubstantiated medical claims, undermining the integrity of civic processes, or manipulating content for false or misleading purposes (whether through generative AI or through deceptive editing).

Civic-process content prohibitions

The Community Guidelines on harmful false information refer to a more detailed Explainer that prohibits content that undermines the integrity of civic processes, deepfake content, or other media manipulated for false or misleading purposes. The Explainer extends the prohibitions to four sub-categories[155]:

Sub-category	Description

[154] Snap Inc., 'Community Guidelines' and 'Harmful False or Deceptive Practices' (updated March 2026).
[155] 'Harmful False or Deceptive Practices', Community Guidelines Explainer Series (updated March 2026).

[redacted]

Procedural interference	Misinformation related to actual election or civic procedures, such as misrepresenting important dates and times or eligibility requirements for participation.
Participation interference	Content that includes intimidation to personal safety or that spreads rumours to deter participation in the electoral or civic process.
Fraudulent or unlawful participation	Content that encourages people to misrepresent themselves to participate in the civic process or to illegally cast or destroy ballots.
Delegitimisation of civic processes	Content aiming to delegitimise democratic institutions on the basis of false or misleading claims about election results, for example.

Sharing such content violates the Community Guidelines irrespective of whether it is AI-generated or user-generated, and irrespective of whether it is generated on Snapchat or on another platform.



[redacted]

Spotlight recommendation of AI-generated content

Following the assessment period, Snap updated its Spotlight recommendation systems to better reward authentic, human-made content. As of July 2026, wholly AI-generated videos are no longer eligible for recommendation on Spotlight. Content that has been enhanced or edited using Snapchat's AI creative tools continues to be eligible for recommendation and includes transparency indicators. Snap's goal is to keep Spotlight a place where authentic creativity has the best opportunity to be discovered, while reducing the risk that wholly AI-generated content is amplified in a way that could blur the distinction between authentic and synthetic content. This change was implemented after the assessment cutoff and is therefore recorded here as a subsequent development relevant to Snap's content-authenticity framework.[156]

Advertising approach

[redacted] As recorded in Section 5.7 [redacted] Advertisements that are not deceptive and otherwise comply with the Advertising Policies are approved to run.

Following the discontinuation of EU political advertising as of October 2025 (consistent with Regulation (EU) 2024/900 and recorded in Section 5.7 and Section 4.3.1), the political-advertising dimension of the advertising approach has reduced operational scope.[157] The political-advertising library remains available for historical disclosure purposes.

User Guidance on Generative AI Features

On-platform features for creating generative AI content are out of scope for DSA purposes (with the limited exception noted above). Notwithstanding the out-of-scope status, Snap maintains a

[156] https://newsroom.snap.com/rewarding-authentic-creativity-on-spotlight

[157] Snap Inc., 'Snap Political and Advocacy Advertising Policies' (accessed 4 August 2026); Regulation (EU) 2024/900 of the European Parliament and of the Council of 13 March 2024 on the transparency and targeting of political advertising, Article 30(2).

[redacted]

generative-AI support site that provides public-facing transparency on what generative AI is and on Snap's practices in respect of generated content.

Snap operates indicators across most Snapchat experiences involving AI to notify users that they are interacting with or viewing AI-generated content. The principal indicators are:

Indicator	Mechanism
Sparkle icon (✨)	Displayed alongside features or outputs powered by generative AI.
Context Cards / disclaimers	Explain when a feature is AI-driven and provide additional context to recipients.
On-screen tooltips and labels	Appear within relevant features (including, by way of example, Dreams and AI Snaps).
Watermarks	Visible watermarks (such as a Snap Ghost with sparkles) are added to exported or saved images created using generative AI tools, in some cases.
Camera Roll watermarks on export	When users export AI-generated images to their Camera Roll, the images will often include a watermark indicating they were created using Snap's AI-powered tools.
Context Cards on shared content	When AI-generated content is shared with others on Snapchat, Snap may include Context Cards to indicate to recipients that the content was created with AI.

Source: Snap Inc., 'Generative AI on Snapchat', Snapchat Support (accessed 4 August 2026).

These signals support user understanding that the relevant content was not created by a human and may not depict real-world events or scenarios. [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

This layered framework supports the mitigation of confusion or misuse risks even where any individual indicator is bypassed. Effective and consistent transparency across the layered framework supports informed user engagement with AI-powered features. During the 2026 assessment period, [redacted]

Enforcement

The enforcement framework for content-authenticity-related violations operates through the integrated framework recorded in Section 5.5 (Enforcement). [redacted]

[redacted]

Partnerships

Tackling content-authenticity-related risks requires (among other things) broad industry-wide technical solutions that have not been definitively identified. Snap is engaged with industry peers and external experts on these questions through the engagements recorded below. The framework operates on the basis that effective content-authenticity measures benefit from cross-industry coordination, technical exchange, and engagement with regulators, civil society and academic networks.

Engagement	Description

[redacted]

[redacted]

[158] Snap Inc., 'Sharing our Response to Civil Society Groups on Election Integrity' (2024); AI Elections Accord, 'The Commitments' (2024).
[159] HackerOne, 'Snap's Safety Efforts with AI Red Teaming from HackerOne' (27 February 2024).

[redacted]

[redacted]

EU AI Act Engagement

Regulation (EU) 2024/1689 (the EU AI Act) entered into force in 2024 and is being implemented through staged entry into application. Snap continues to monitor and assess the staged implementation of the EU AI Act and its implications for Snap's products and services. Relevant areas include the Article 50 transparency obligations for AI-generated content, the General-Purpose AI Code of Practice, the development of harmonised standards under Article 40, and related voluntary initiatives such as the [redacted]

DSA Code of Conduct on Disinformation

The Code of Practice on Disinformation (originally adopted in 2018 and strengthened in 2022) transitioned to operate under the DSA framework as the DSA Code of Conduct on Disinformation, with operational status under the DSA framework from 1 July 2025. The Code addresses commitments across multiple dimensions bearing on content authenticity, including in respect of demonetisation of disinformation, transparency of political advertising, integrity of services (including bot detection), empowering users (with media literacy and content evaluation tools), empowering the research community, and empowering the fact-checking community. Snap was not a signatory to the DSA Code of Conduct on Disinformation during the 2026 assessment period. Accordingly, this Report does not rely on adherence to the Code as a mitigation or represent that Snap is subject to its signatory reporting and auditing framework.[160]

[160] European Commission, 'Codes of Conduct under the Digital Services Act' (updated 2 July 2026).

The substantive content-authenticity framework (the two-track Creation/Dissemination architecture; the internal Generative AI Policies; the Community Guidelines applicability to AI-generated content; the multi-layered user-guidance framework; the integrated enforcement framework recorded in Section 5.5) is unchanged in its core architecture.

Conclusion

Content authenticity is a challenging topic without a single technical solution. Snap operates a multi-layered framework that combines internal policies governing the creation of generative-AI content; the Community Guidelines and Terms applicability to AI-generated content as a content-format-agnostic baseline; user-facing indicators across most AI-engagement experiences; the integrated enforcement framework; and continuing engagement with industry partners, the European Commission, the AI Office, Member State authorities, and external researchers.

The framework reflects measures identified in the Commission's Guidelines on the Mitigation of Systemic Risks for Electoral Processes and operates alongside the substantive Category 1, Category 3 and Category 4 mitigation sets recorded across Section 4. Snap will continue to monitor developments and industry practice on a continuing basis, including in respect of the appropriate use of prominent markings and other measures to distinguish content that falsely appears to be authentic, and will continue to engage with the regulatory framework as the EU AI Act, the DSA Code of Conduct on Disinformation, and related instruments are operationalised.

The mitigations recorded in this Section, in combination with the other mitigations recorded across Section 5, support the conclusions recorded across the underlying Category 1, Category 3 and Category 4 assessments that Snap's mitigations for the systemic risks identified in respect of Snapchat's in-scope services are reasonable, proportionate and effective.

5.10 Trusted Flaggers

Introduction

Section 5.10 sets out Snap's Trusted Flagger Program, the operational framework through which qualified third-party organisations submit notices of illegal content and Community Guidelines violations to Snap on a priority-handling basis. The Program operates as a key complement to the Article 16 DSA notice-and-action framework recorded in Section 5.4.6 and to the substantive moderation framework recorded in Section 5.4.

The Snap Trusted Flagger Program was established to enable non-profit organisations, non-governmental organisations (NGOs), select government agencies, and safety partners to support the Snapchat community by providing a dedicated, prioritised reporting channel for content that violates Snap's Community Guidelines or constitutes illegal content. The Program

[redacted]

has been in operation for a number of years and has been developed iteratively in response to evolving systemic risks and the corresponding mitigation framework recorded across Section 4 of this Report.

With the entry into force of the Digital Services Act, Snap's Trusted Flagger Program was adapted to incorporate the Article 22 DSA framework. The Program now operates two complementary tracks: trusted flaggers awarded that status by a Digital Services Coordinator under Article 22(2) DSA, and trusted flaggers who have entered into a direct agreement with Snap. Both tracks are subject to the same substantive operational framework recorded below.

Several trusted flagger relationships support the Snap mitigation framework in relation to substantive risk categories:

- [redacted]

Two-track structure

Snap receives reports from trusted flaggers operating under two complementary tracks:

Track	Description
DSA-appointed trusted flaggers	Entities awarded trusted flagger status by a Member State Digital Services Coordinator under Article 22(2) DSA and listed on the European Commission's public register under Article 22(5) DSA. Snap monitors the Commission's published list on a continuing basis and onboards relevant entities upon publication. The substantive operational framework (webform reporting, priority handling and outcome notification) applies uniformly to DSA-appointed trusted flaggers.

Snap direct-agreement trusted flaggers	Entities that have entered into a direct agreement with Snap to operate as a trusted flagger outside the Article 22 DSA framework, typically pre-dating or independent of the DSA's coming into force. These relationships continue to support the substantive framework and are evaluated against criteria including geographic coverage, area of expertise, anticipated volume of reports, and demonstrated effectiveness. The substantive operational framework applies uniformly to Snap direct-agreement trusted flaggers.

Article 22 DSA framework

Snap's operational implementation of Article 22 DSA is mapped against the relevant sub-paragraphs of the Article below:

Article 22 DSA element	Operational implementation
Article 22(1) DSA	Notices submitted by trusted flaggers acting within their designated area of expertise are processed and decided upon with priority and without undue delay. The Snap Trusted Flagger webform feeds into a high-priority review channel, with reports relating to the most serious harms prioritised for prompt review.
Article 22(2) DSA	Trusted flagger status is awarded by the Digital Services Coordinator of the Member State where the applicant is established, on application by an entity that demonstrates particular expertise and competence for detecting, identifying and notifying illegal content; that is independent from any provider of online platforms; and that carries out its activities for the purposes of submitting notices diligently, accurately, and objectively.
Article 22(3) DSA	Trusted flaggers publish, at least once a year, easily comprehensible and detailed reports on notices submitted in accordance with Article 16 DSA during the relevant period. The Snap framework supports the data inputs necessary for trusted flaggers to fulfil their Article 22(3) reporting obligations.
Article 22(5) DSA	The European Commission publishes the list of awarded trusted flaggers on a publicly accessible website, kept updated. Snap monitors the Commission's published list (referenced below) on a continuing basis and follows the standard onboarding process upon notice of new entries relevant to Snapchat's in-scope services.

[redacted]

Article 22(6) DSA	Where a provider of an online platform has information indicating that a trusted flagger submitted a significant number of insufficiently precise, inaccurate or inadequately substantiated notices, the provider may communicate this information to the Digital Services Coordinator that awarded trusted flagger status. The Snap framework includes operational triggers for such communication.
Article 22(7) DSA	Where the Digital Services Coordinator confirms that the trusted flagger conditions are no longer met, it may suspend or revoke trusted flagger status. The Snap framework recognises and follows DSC determinations on trusted flagger status.

By the end of the 2026 assessment period, the Article 22 trusted-flagger framework had continued to expand. On 29 May 2026, the European Commission reported that more than 70 trusted flaggers had been designated across the European Union, and the Commission's public Article 22(5) register was subsequently updated on 8 June 2026[161]. Snap monitors that register on a continuing basis and onboards relevant entities through the standard onboarding process described below.

Onboarding

DSA-appointed trusted flaggers

Onboarding of DSA-appointed trusted flaggers is initiated upon publication of the appointment on the European Commission's public register under Article 22(5) DSA. Snap monitors the register on a continuing basis. Upon notice of a new appointment, Snap reviews whether the trusted flagger's mission area is relevant to Snapchat's in-scope services and, where it is, establishes communication using the contact email supplied by the Member State Digital Services Coordinator. Where the appointment is relevant, the standard onboarding process recorded below is initiated.

Snap direct-agreement trusted flaggers

Snap considers applications from organisations seeking to operate as trusted flaggers under direct agreement, outside the Article 22 DSA framework. The factors considered in evaluating an application include:

- [redacted]

[161] European Commission, "Commission seeks feedback on draft trusted flaggers guidelines under the Digital Services Act" (29 May 2026); European Commission, "Trusted flaggers under the Digital Services Act (DSA)" (register updated 8 June 2026).

[redacted]

- [redacted]
- [redacted]
- [redacted]

Onboarding package

Once a trusted flagger is appointed (DSA track) or accepted (direct-agreement track), Snap provides the following onboarding package:



Reporting and review process

Webform-based reporting

Trusted flaggers submit reports through a dedicated, confidential webform specific to the Trusted Flagger Program. The webform is not linked from any public-facing surface and is accessible only to trusted flaggers who have been provided with the direct link as part of the onboarding package. Webform reports feed into a high-priority review channel within Snap's Trust & Safety team.

[redacted]

Priority review

Snap processes Trusted Flagger reports through a high-priority review channel, consistent with the Article 22(1) DSA requirement that qualifying notices be processed and decided upon with priority and without undue delay. Reports relating to the most serious harms are prioritised and reviewed as promptly as practicable.

Review and outcome

Snap's Trust & Safety team reviews each report and takes any appropriate enforcement action, in line with the framework recorded in Section 5.5 (Enforcement). Where additional information is required to enable full investigation, the Trust & Safety team requests this from the trusted flagger. Once a decision has been made, Snap notifies the trusted flagger of the outcome of the review, including whether enforcement action has been taken in relation to the reported content or accounts.

The webform-based pathway is designed to supplement, rather than replace, the in-app reporting pathway available to all Snapchat users. In-app reporting remains a core component of the broader notice-and-action framework recorded in Section 5.4.6.

Coordination with Digital Services Coordinators

Where Snap's Trust & Safety team identifies trends impacting the quality of reports submitted by a particular trusted flagger - [redacted] - Snap may communicate this information with the Digital Services Coordinator that awarded the trusted flagger status, with the objective of identifying a resolution. The framework operates on a continuing basis and supports the iterative improvement of report quality.

Trusted Flagger Program Trends

EU Trusted Flagger reports overview

During the 2026 assessment period, Snap received [redacted] unique reports through the audited EU Trusted Flagger reporting pathway, comprising [redacted] reports in H2 2025 and [redacted] reports in H1 2026. The H1 2026 volume therefore increased by approximately [redacted] relative to H2 2025.

For purposes of the aggregate report volumes in this Section, each [redacted] ticket is counted once, irrespective of whether the ticket carries more than one relevant internal e-safety-10x classification tag. The underlying audit dataset contains [redacted] tag-level records corresponding to [redacted] unique [redacted] tickets. The unique-ticket measure is used here because it reflects the number of distinct reports received rather than the number of internal classifications attached to those reports.

[redacted]

The audit population covers the period from 1 July 2025 to 30 June 2026 and includes reports identified through the relevant Trusted Flagger reporting tags. Non-EU requester accounts were excluded from this assessment-period dataset. The figures in this subsection should therefore be understood as EU assessment-period reporting data and should not be combined with the broader historical figures that included direct-agreement trusted flaggers outside the European Union.

The increase in report volume is presented as an observed operational trend. The available data does not, by itself, establish the cause of the increase, and Snap has therefore not attributed the change to any particular product, enforcement or reporting-flow development.

Country-level breakdown

Reports are attributed to the country of the originating Trusted Flagger, rather than to the country of the user or content that is the subject of the report. The assessment-period unique-ticket figures are:


[redacted]

Reporting remained concentrated in a relatively small number of Member States. [redacted] reports, followed by [redacted]. This concentration reflects the reporting activity of established Trusted Flagger organisations in those Member States and should not be interpreted as indicating a corresponding geographic concentration of the underlying harms.

2026 assessment-period Trusted Flagger detail

[redacted]

The table below records the unique-ticket report volumes by originating Trusted Flagger organisation for H2 2025 and H1 2026. Where an organisation used more than one reporting address during the period, the relevant requester accounts have been aggregated at organisation level. Each [redacted] ticket is counted once.

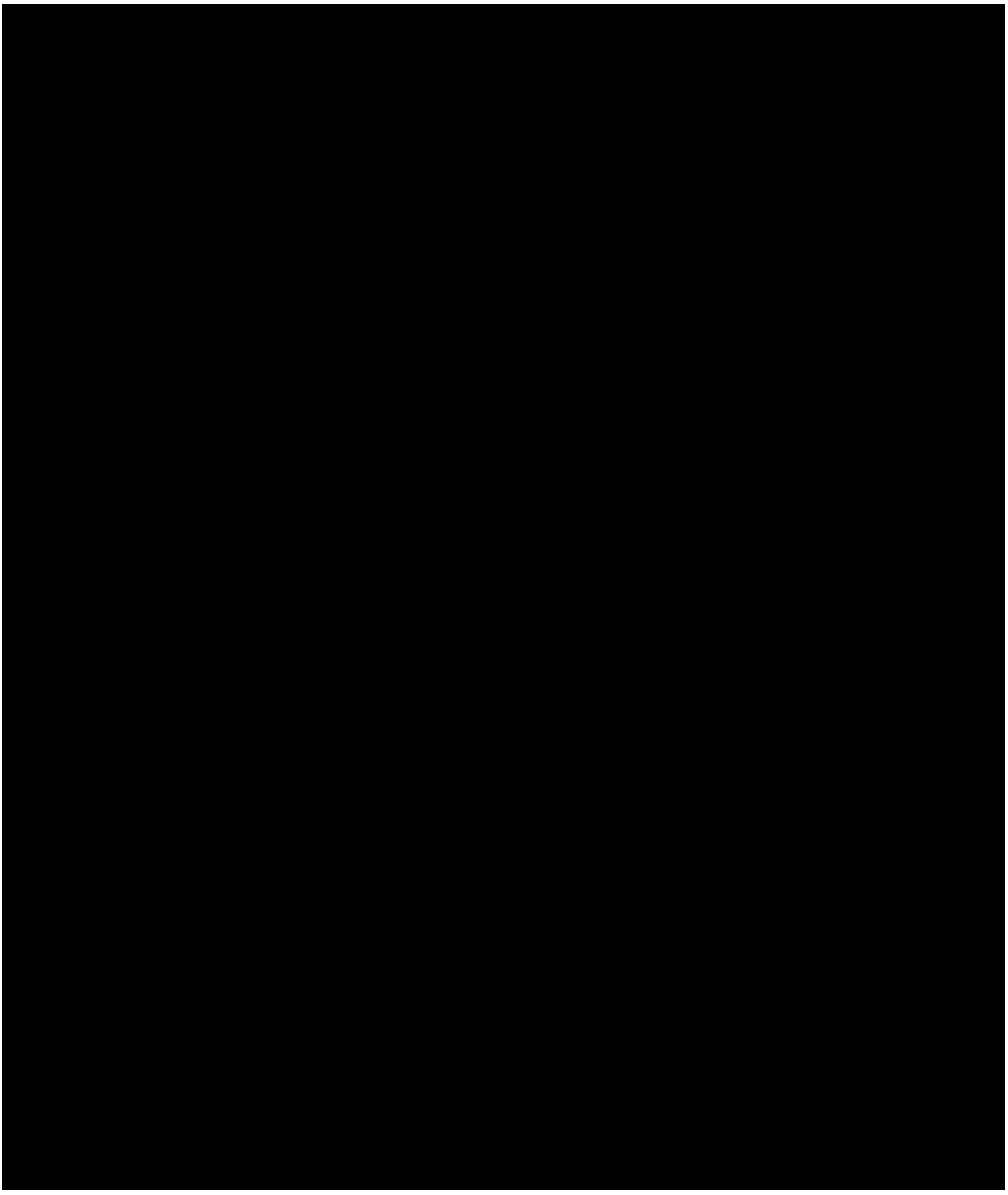
[redacted]

Report volumes were particularly concentrated in [redacted] unique reports in H2 2025 and [redacted] in H1 2026, representing approximately three quarters of the EU Trusted Flagger reports received in each half-year. The table reports activity, rather than the full population of organisations eligible or onboarded to submit Trusted Flagger notices; an organisation with no report during the assessment period will therefore not appear.

Cross-references by mission area

[redacted]

The substantive operational coordination between trusted flagger relationships and the Section 4 risk-assessment framework is mapped below. The cross-references are recorded in detail in the relevant Section 4 assessments; the table below provides a summary view for ease of reference.



[redacted]

Coordination with the Article 23 DSA misuse-suspension framework

The Trusted Flagger Program operates alongside the Article 23 DSA misuse-suspension framework recorded in Section 5.5 (Protections against Misuse). The two frameworks address different operational concerns:

- **Article 22(6) DSA.** Addresses patterns of insufficiently precise, inaccurate or inadequately substantiated trusted flagger notices, with the operational pathway being communication to the relevant Digital Services Coordinator with a view to resolution. This is the principal mechanism for managing report-quality concerns within the Trusted Flagger framework.
- **Article 23 DSA.** Addresses patterns of bad-faith reporting (manifestly unfounded notices and complaints) by individual users, with the operational pathway being suspension of the processing of notices and complaints from such users for up to one year. The framework is recorded in Section 5.5.

The two frameworks operate as complementary protections against report-system misuse and are administered through coordinated procedures within the Trust & Safety team.

The substantive Trusted Flagger framework, including the two-track structure, priority-review arrangements, Article 22 implementation and coordination with Digital Services Coordinators on report-quality concerns, remains unchanged in its core architecture.

Conclusion

Snap operates a carefully managed Trusted Flagger Program, including an Article 22 DSA track for designated EU Trusted Flaggers and a complementary direct-agreement track, covering substantive harm categories identified across Section 4 of this Report. Both tracks operate within the substantive operational framework described above. During the 2026 assessment period, the Program continued to evolve through the transition to webform-based reporting, refreshed onboarding materials and the continued expansion of the Article 22 Trusted Flagger network. The audited EU reporting pathway recorded [redacted] unique reports during the assessment period, with report volume increasing from [redacted] in H2 2025 to [redacted] in H1 2026.

As recorded across Section 4, Snap has concluded that the Trusted Flagger Program, in combination with the other mitigations recorded across Section 5, is a reasonable, proportionate and effective mitigation measure for the systemic risks identified for Snapchat's in-scope services. The Program will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes, with particular attention to: the operational outputs of the 2026 webform transition; the integration of newly-appointed DSA Trusted Flaggers into the existing framework; and the continued operational coordination with Digital Services Coordinators on report-quality concerns and broader Article 22 DSA matters.

5.11 Dispute Settlement Bodies

Introduction

Section 5.11 sets out Snap's operational framework for engagement with out-of-court dispute settlement ("ODS") bodies certified under Article 21 DSA. The framework operates as the external dispute-resolution pathway available to recipients of the Snapchat service who are dissatisfied with the outcome of Snap's internal complaint-handling process recorded in Section 5.5 (Enforcement). Snap's community support teams operate the internal complaint-handling pathway as the principal first-instance resource for resolving queries and complaints; the Article 21 DSA framework operates as the external escalation pathway recognised by the Regulation.

Consistent with Article 21 DSA, recipients of the Snapchat service are informed of their right to seek out-of-court dispute resolution where they remain dissatisfied with the outcome of Snap's internal complaint-handling process. The substantive operational implementation of this framework is recorded below, together with operational data for H1 2026 and Snap's observations on the broader operation of the framework as it has developed since the entry into force of the DSA.

Article 21 DSA framework

The substantive operational implementation of Article 21 DSA is mapped against the relevant sub-paragraphs of the Article below:

Article 21 DSA element	Operational implementation
Article 21(1) DSA	Recipients of the service, including individuals or entities that submitted notices, may select any out-of-court dispute settlement body certified under Article 21(3) DSA to resolve disputes relating to decisions referred to in Article 20(1), including complaints not resolved through the internal complaint-handling system. The selected body's certified expertise and language coverage must encompass the dispute. Snap's notice-and-appeal framework recorded in Section 5.5 operates as the internal complaint-handling pathway associated with the Article 21 ODS framework.
Article 21(1) DSA (continued)	Recipients are entitled to select any out-of-court dispute settlement body certified under Article 21(3) DSA whose certified area of expertise and language coverage encompass the relevant dispute. Snap monitors the Commission's overview of certified ODS bodies on a continuing basis.

Article 21(2) DSA	Online platforms engage, in good faith, with the selected certified ODS body with a view to resolving the dispute. Snap's operational engagement with certified ODS bodies follows this framework, recorded in detail below.
Article 21(3) DSA	ODS body certification by the relevant Digital Services Coordinator depends on the body satisfying the conditions of impartiality, independence, expertise, accessibility, fairness and procedural soundness. Snap recognises and follows DSC certification determinations and engages with bodies whose certified areas of competence cover disputes relating to Snapchat.

Certified ODS bodies status

As of 2 July 2026, shortly after the close of the assessment period, eleven out-of-court dispute settlement bodies had been certified by Member State Digital Services Coordinators under Article 21(3) DSA. Recipients of the Snapchat service may refer their case to any certified body whose area of expertise and language coverage encompass the relevant dispute. Snap monitors the European Commission's published list on a continuing basis.[162]

Engagement process

Where a recipient of the Snapchat service submits a dispute to a certified ODS body, the process operates as follows:

- **Recipient submission.** The recipient submits the dispute directly to the selected certified ODS body, which conducts its own intake and admissibility screening.
- **Body-to-Snap contact.** The certified ODS body contacts Snap via the dedicated webform to initiate the dispute settlement process. The webform is the standardised intake mechanism for Article 21 DSA matters and supports consistent ingestion into Snap's internal coordination workflow.
- **Good-faith engagement.** Consistent with Article 21(2) DSA, Snap engages with the certified body in good faith, with a view to resolving the dispute, following Snap's established policies and procedures (including the substantive Terms framework recorded in Section 5.2 and the enforcement framework recorded in Section 5.5).
- **Resolution and decision.** The matter is resolved either by settlement reached during the engagement, or by a decision issued by the certified ODS body. Snap's operational engagement records the outcome and any consequential platform action.

[162] European Commission, 'Out-of-court dispute settlement bodies under the Digital Services Act (DSA)' (register last updated 2 July 2026).

[redacted]

The 2026 transition from email-based contact (the legacy dsa-enquiries@snapchat.com pathway) to webform-based contact is consistent with the broader 2026 operational shift toward webform-based intake across multiple stakeholder-facing pathways (recorded in Section 5.10 in respect of trusted flagger reporting). The webform pathway supports more consistent ingestion and tracking of Article 21 matters within Snap's internal coordination workflow.

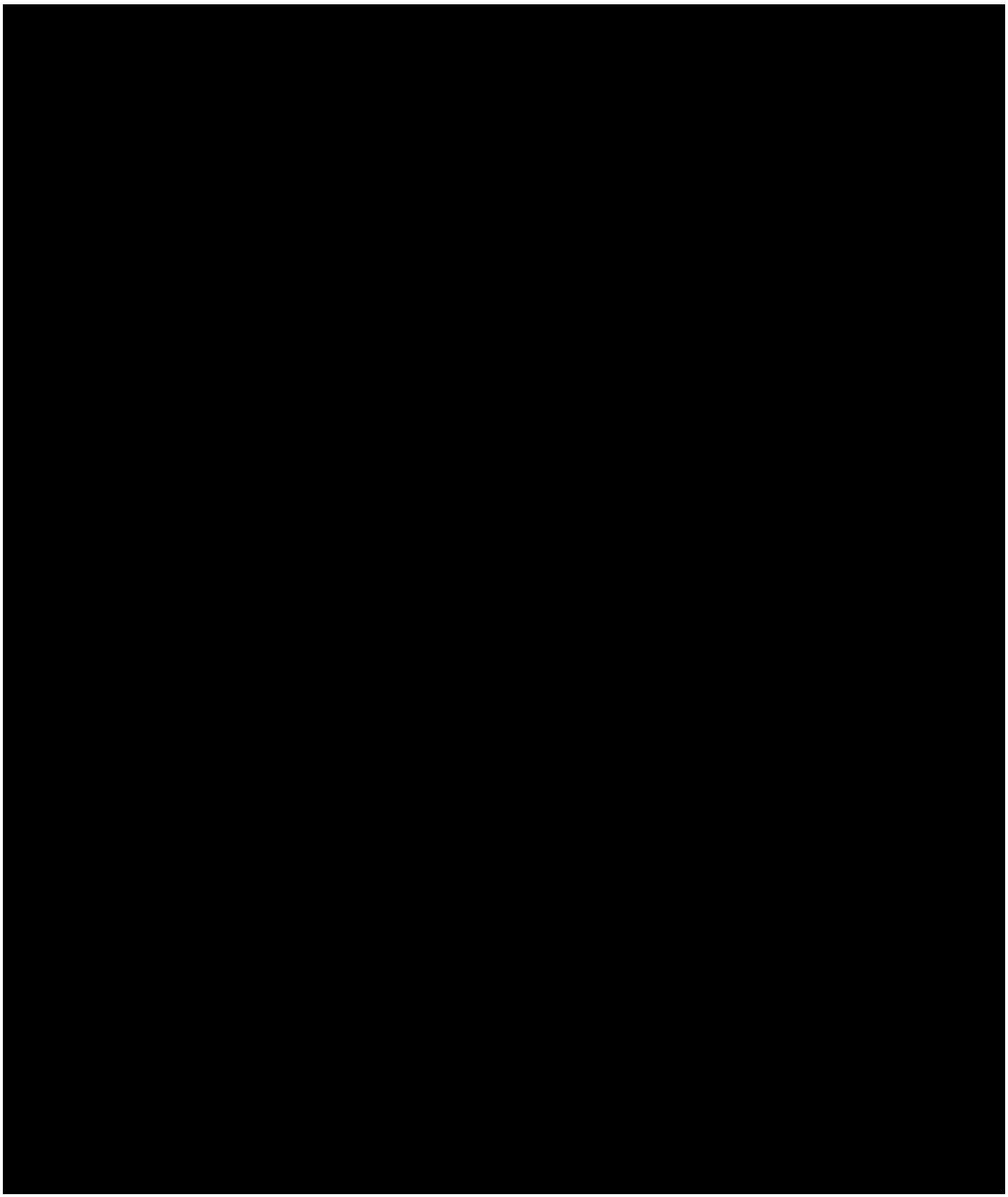
Enquiries

[redacted]

Data relating to disputes received between 1 July and 31 December 2025 can be found in Snap's H2 2025 EU DSA Transparency Report.

Stakeholder observations on the operation of the framework

[redacted]



The substantive Article 21 DSA framework (good-faith engagement with certified bodies, alignment with the internal complaint-handling framework recorded in Section 5.5, the recommendations on framework evolution recorded above) is unchanged in its core architecture.

Conclusion

Snap remains committed to resolving user disputes effectively and consistently with Article 21 DSA. The framework operates through the internal complaint-handling pathway recorded in Section 5.5 as the principal first-instance resource, with the Article 21 DSA ODS pathway operating as the external escalation tier. Snap engages in good faith with certified ODS bodies whose certified areas of competence cover disputes relating to Snapchat and continues to support the effective evolution of the Article 21 framework, including through the operational observations recorded above.

As recorded across Section 4, Snap has concluded that the Article 21 DSA dispute-resolution framework, in combination with the other mitigations recorded across Section 5, is reasonable, proportionate and effective for the systemic risks identified for Snapchat's in-scope services. The framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes following the 2026 assessment period, with particular attention to: the operational outputs of the 2026 webform transition for ODS body engagement; the continued development of the certified-body landscape; and the operational implementation of any structural evolution of the framework at the EU level.

5.12 Codes and Crisis Protocols

Introduction

Section 5.12 sets out three complementary frameworks supporting Snap's mitigation architecture: (i) industry cooperation through international fora and trade associations; (ii) participation in voluntary codes of conduct under Article 45 DSA; and (iii) crisis protocols under Article 48 DSA, supported by Snap's broader content crisis management framework. Each framework operates as a structural complement to the substantive mitigations recorded across Sections 5.1–5.11, with cross-references to the Section 4 risk assessments that relate to the relevant cooperation, code or crisis-response architecture.

Cooperation with external stakeholders is a meaningful element of Snap's risk-mitigation framework. Knowledge-sharing and best-practice development with experts and peers strengthen the effectiveness of Snap's internal risk-mitigation measures, in particular in respect of the substantive harm categories where cross-platform signal-sharing or cross-industry calibration is operationally important (CSAM, hate speech, terrorism and violent extremism, cross-platform sextortion, suicide and self-harm content, and minor protection more broadly). The cooperation framework recorded below operates alongside the substantive trusted-flagger relationships

recorded in Section 5.10 and the substantive moderation and enforcement frameworks recorded in Sections 5.4 and 5.5.

Industry cooperation

Snap participates in a structured set of industry fora, trade associations and standards-development organisations, each engaging different elements of the substantive mitigation framework recorded across this Report. The principal current memberships and engagements are recorded below:

Forum / partnership	Description
EU Internet Forum (EUIF)	Active member and contributor. The EUIF provides a collaborative environment for EU governments, the internet industry, and other experts and partners to discuss and address the challenges posed by the presence of malicious and illegal content online - originally focused on terrorism and violent extremism and on grooming for child sexual exploitation and abuse, with the scope expanded in recent years to address drug sales online and human trafficking. Snap participates in the EUIF working groups relevant to the substantive risk categories assessed in this Report.
Tech Coalition	Member. The Tech Coalition is an alliance of global technology companies working to combat child sexual exploitation and abuse online. The Tech Coalition coordinates industry's overall effort to combat child sexual abuse online, providing resources, education and capacity-building to technology companies, and serving as a resource for external stakeholders - from policy-makers to journalists - on industry-level responses. Snap participates in the Tech Coalition's Working Group on Generative AI Content and the GenAI Briefing Subgroup (recorded in Section 5.9).
WeProtect Global Alliance	Board member. WeProtect Global Alliance brings together the private sector, government and civil society to drive change to help protect children from sexual abuse online. Snap's Board membership reflects its substantive commitment to the cross-platform child-protection framework.
Alliance to Better Protect Minors Online	Former member - initiative discontinued in July 2024. The Alliance was a self-regulatory initiative aimed at improving the online environment for children and young people through structured stakeholder engagement and best-practice exchange. Snap was a

	member from inception through to the discontinuation of the initiative in July 2024.
CIPL (Center for Information Policy Leadership)	Member. CIPL is a global privacy and data policy think and do tank based in Washington, DC, Brussels and London. Snap's engagement with CIPL covers global solutions and best practices for privacy and responsible data use, including in respect of teenagers and young adults; CIPL's AI work (including the dedicated AI Project and the Brazilian AI Project) is referenced in Section 5.9.
Future of Privacy Forum (FPF)	Member. FPF is a non-profit organisation supporting privacy leadership and scholarship. Snap's engagement with FPF covers a range of matters including augmented reality, artificial intelligence, biometric data, and children's rights. FPF's AI Governance work and responsible Generative AI initiatives are referenced in Section 5.9.
CERRE (Center on Regulation in Europe)	Member. CERRE is a not-for-profit Brussels-based think tank supporting policy-oriented research and engagement on regulation in Europe and beyond. Snap's engagement with CERRE covers reports and issue papers on regulation challenges in the technology, media and telecommunications sector.
Internet Watch Foundation (IWF)	Member. IWF is an independent, non-profit organisation focused on preventing child sexual abuse online. Snap sits on the IWF Funding Council, and Snap's Head of Public Policy, UK & Ireland, is an industry trustee on the IWF Board. The IWF's substantive work on CSAM detection is recorded in Section 4.1.1.
Thrive	Member. Thrive is an industry signal- and best-practice-sharing initiative focused on suicide and self-harm content. The substantive engagement is recorded in Section 4.1.8.
DOT Europe	Member. DOT Europe coordinates EU-level engagement on privacy, security, safety, and content policy issues across the technology industry.
ITI (Information Technology Industry Council)	Member. ITI coordinates EU-level engagement on privacy, security and safety policy issues; Snap's participation in ITI's AI Futures Initiative is recorded in Section 5.9.

Codes of Practice (Article 45 DSA)

Article 45 DSA framework

Article 45 DSA establishes that the European Commission and the European Board for Digital Services encourage and facilitate the drawing up of voluntary codes of conduct at Union level to contribute to the proper application of the DSA. Codes adopted under this framework are required to set out specific objectives, contain key performance indicators to measure achievement, and are subject to monitoring and evaluation by the Commission and the Board. The Commission and the Board publish their conclusions on whether the codes meet the aims of the DSA in addressing illegal content categories and systemic risks.

The framework remains in active development; transitional questions about how existing pre-DSA codes are assessed against DSA requirements continue to be worked through. Snap supports industry-wide efforts to promote risk-mitigation practices through voluntary codes, recognising that the self-regulatory approach contemplated by Article 45 supports the development of flexible, proportionate and risk-based solutions to industry-wide problems. The framework operates effectively where it engages all relevant industry players - regardless of size and resources - and where the resulting codes are calibrated to the operational realities of platforms with diverse business models, user populations, and risk profiles. Snap's continuing engagement with the Article 45 framework reflects this position.

Snap's current code participation status

Code	Status	Detail
EU Code of Conduct on Countering Illegal Hate Speech Online (revised 20 January 2025)	Signatory	Signed onto the original Code in 2018; complying with the revised DSA-aligned Code as part of Snap's DSA compliance programme since 20 January 2025.
FSM Code of Conduct (Freiwillige Selbstkontrolle Multimedia-Diensteanbieter, Germany)	Signatory	Joined FSM in September 2017 as a full member; the FSM Code is a binding voluntary self-regulation instrument under the German Interstate Treaty on the Protection of Minors in the Media (JMStV).
EU Code of Conduct on Disinformation (operative under the DSA framework as of 1 July 2025)	Not a signatory	Snap was not a signatory to the DSA Code of Conduct on Disinformation during the 2026 assessment period. Accordingly, this Report does not rely on adherence to the Code as a mitigation or represent that Snap is subject to its signatory reporting and auditing framework.
EU Code of Conduct for Age Appropriate Design (AAD Code)	Special-group work concluded	Snap participated in the development process for the proposed Code. The European Commission has since stated that the work of the special group established to assist with drafting the Code has

[redacted]

		ended. This Report therefore does not present the proposed Code as an operative Article 45 DSA code or as an active 2026 mitigation. Snap's current operational measures are mapped against the Commission's Article 28 DSA Guidelines in Section 5.8.

EU Code of Conduct on Countering Illegal Hate Speech Online

Snap signed onto the EU Code of Conduct on Countering Illegal Hate Speech Online in 2018, two years after the Code was first introduced in 2016. The Code has, over the period of its operation, become a meaningful instrument for cooperation among industry, civil-society organisations (CSOs), and public authorities on the substantive challenge of online hate speech. The framework's core mechanism - the regular monitoring exercises conducted in collaboration with a network of specialised CSOs across EU Member States - has supported the development of effective notice-and-action mechanisms, structured platform/CSO partnerships, and best-practice training on countering hateful rhetoric and prejudice.

[redacted]

Annex (December 2022) and DSA alignment (January 2025)

In December 2022, Snap participated alongside other industry signatories in the publication of an Annex to the existing Code, strengthening the framework on cooperation between IT companies and CSOs beyond the remit of the monitoring exercises. [redacted] The DSA-aligned revision was finalised on 20

[163] European Commission, 'First results published under the revised Code of Conduct on Countering Illegal Hate Speech Online+' (10 April 2026).

[164] European Commission, 2025 Monitoring Exercise under the Code of Conduct on Countering Illegal Hate Speech Online+; Snap Inc., Snapchat Report under the Code of Conduct on Countering Illegal Hate Speech Online+ (2026).

[redacted]

January 2025; since that date, Snap has complied with the revised Code as part of its DSA compliance programme.

FSM Code of Conduct

In September 2017, Snap joined the Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM), an officially recognised voluntary self-regulation association for the protection of minors in online media in Germany. As a full member, Snap has signed the FSM Code of Conduct, which aims to:

- protect children and adolescents from content offered on digital services that could endanger or impair their development;
- provide protection from content that violates the right to human dignity or other aspects of legal protection under the German Interstate Treaty on the Protection of Minors in the Media (JMStV);
- prevent the depiction of CSAM on participating services;
- ensure the distribution of age-appropriate content and protect children and adolescents from developmentally harmful content;
- provide reporting systems for user complaints; and
- prevent the distribution of advertising on participating services that violates the JMStV.

Once signed, the FSM Code is binding on members and violations are subject to the framework's sanction provisions. Snap has not violated the FSM Code to date and remains in continuing dialogue with FSM. The FSM membership operates alongside the Article 22 DSA trusted flagger relationship with the FSM organisation recorded in Section 5.10.

[redacted]

[redacted]

EU Code of Conduct for Age Appropriate Design

Snap participated in the development process for the proposed EU Code of Conduct for Age Appropriate Design. The Commission has since confirmed that the work of the special group established to assist with drafting the Code has ended. Accordingly, this Report does not treat the proposed Code as an operative Article 45 DSA instrument or an active 2026 workstream. Snap's measures for the protection of minors are assessed against the Commission's Article 28 DSA Guidelines in Section 5.8.

Article 28 DSA Guidelines

In the period since publication of the European Commission's Guidelines on Article 28 DSA in 2025, Snap has assessed the operational position of Snapchat's in-scope services against the Guidelines' principal recommendations. Section 5.8 (Protection of Minors) maps Snap's measures against the principal sections of the Guidelines.

Crisis Response and Crisis Protocols (Articles 36 and 48 DSA)

Snap operates a structured set of crisis-management protocols supporting swift response to unexpected incidents, with the objective of minimising impact on the service, the user community, and Snap's operations. The crisis-management framework operates across three principal domains:

- **Security Incident Response Policy (SIRP).** Provides the framework for security incident response operations across Snap products and services, company information, and company systems (including company information and systems managed by third parties). The SIRP operates as the principal infrastructure-level crisis-management instrument.
- **Privacy incident protocol.** Provides the framework for response to privacy incidents, including potential data breaches. The protocol operates alongside the substantive privacy and data-protection framework recorded in Section 4.2.4.
- **Content Crisis Management and Response Protocol.** Provides the framework for response to content-level crisis events affecting Snapchat's in-scope services. The protocol operates as Snap's principal internal content-crisis management framework and is described in detail below.

Articles 36 and 48 DSA framework

Article	Operational implementation
Article 36 DSA - Crisis response mechanism	Authorises the Commission, in the event of a crisis, to adopt decisions requiring providers of VLOPs to take specific actions to limit any contribution to the crisis. Crisis is defined as extraordinary circumstances leading to a serious threat to public security or public health in the EU or in significant parts of it.

[redacted]

Article 48 DSA - Crisis protocols	Provides for the establishment of crisis protocols on a voluntary basis at Union level, addressing extraordinary circumstances affecting public security or public health. Snap's Content Crisis Management and Response Protocol is an internal crisis-management framework designed to support readiness for crisis situations and to facilitate compliance with any applicable Commission decision under Article 36 DSA or any Article 48 crisis protocol in which Snap participates. It is not itself presented as a crisis protocol established under Article 48 DSA.
Article 35(1)(h) DSA	Requires VLOPs, in respect of identified systemic risks, to consider initiating or adjusting cooperation with other providers through the codes of conduct and crisis protocols referred to in Articles 45 and 48. The substantive cross-references to Article 35(1)(h) in this context are recorded across the Section 4 risk-assessments.

Content Crisis Management and Response Protocol

In many cases, Snap's substantive content moderation policies form part of Snap's internal response to unexpected external events that may affect the quality of content surfacing on Snapchat's in-scope services. [redacted]

Routine escalation pathway

[redacted]

[redacted]

Conclusion

Snap's framework for industry cooperation and voluntary codes of conduct under Article 45 DSA, together with its preparedness for the Article 36 DSA crisis-response mechanism and participation in any crisis protocol developed under Article 48 DSA, constitutes the principal architecture for cross-platform and cross-stakeholder risk-mitigation work supporting the substantive measures recorded across Sections 5.1–5.11. The framework reflects Snap's continuing commitment to constructive engagement with the European Commission, Member State authorities, civil-society organisations, peer platforms, and standards-development organisations on the substantive challenges of online safety, with prioritisation calibrated to the substantive risk profile recorded in Section 4 of this Report.

As recorded across Section 4, Snap has concluded that the framework recorded in this Section, in combination with the other mitigations recorded across Section 5, is reasonable, proportionate and effective for the systemic risks identified for Snapchat's in-scope services. The framework will continue to be subject to ongoing monitoring through the Section 6 risk-detection processes during the 2026 assessment period, with particular attention to: the operational outputs of any 2026 changes to the cooperation portfolio; and any 2025–2026 internal crisis-management activations that supplement the operational case study record.

6. Ongoing Risk Detection and Management

Article 35(1)(f) DSA identifies, as a category of mitigation measure, the reinforcement of internal processes, resources, testing, documentation and supervision of a platform's activities, in particular as regards the detection of systemic risk. Article 42(4)(b) DSA requires reporting on the specific mitigation measures put in place pursuant to Article 35(1), and Article 42(4)(e) requires reporting on the results of audits, the independent compliance function and the supervisory fee. This Section 6 records the ongoing risk detection and management framework through which Snap operates and discharges those obligations.

The framework comprises four principal components, preserved from prior Reports:

- **Platform Principles-based Framework.** A values- and rights-based framework used to identify, investigate and mitigate risk across Snapchat's in-scope services. The framework draws on international human rights principles, digital-policy best practice, and risk-based metrics including prevalence and severity analyses. See Section 6.1.
- **Cross-functional governance.** The designation of a senior, cross-functional team responsible for applying the framework and assessing its outcomes, including the DSA Independent Compliance Function and the DSA Cross-Functional Governance Team. See Section 6.2.
- **Internal resources supporting risk detection.** A repository of internal resources supporting the detection and management of risk, including harm severity assessments, prevalence metrics (Policy-Violating Prevalence and Violative View Rate), reporting-data reviews, the Digital and Data Impact Assessment (DDIA) framework, and a library of Terms and policy resources. See Sections 6.3 and 6.4.
- **Continuing improvement and external assurance.** Ongoing measurement and improvement through the Digital Well-Being Index initiative, the Safety Advisory Board, the Snap Council for Digital Well-Being, dedicated DSA Risk Assessment workshops, and the Article 37 independent audit. See Sections 6.6, 6.7 and 6.8.

Relationship with Section 5 (Specific Mitigations)

Section 5 describes the specific mitigation measures put in place pursuant to Article 35(1)(a)-(e) and (g)-(k) DSA, organised by mitigation category. This Section 6 describes the internal assurance and ongoing-detection processes through which the operation of those mitigations is monitored, evaluated and adjusted over time, corresponding to Article 35(1)(f). The two Sections are designed to be read in combination: Section 5 records what Snap does; Section 6 records how Snap knows whether what it does is working, and how it adjusts over time.

6.1 Platform Principles-based Framework

Introduction

[redacted]

[redacted]

[redacted]

[redacted]



Ongoing engagement with subject-matter experts continued in 2026 through the [redacted] These engagement channels are described in Section 6.7.

Framework structure

The Framework comprises two complementary parts, each borrowing from longstanding [redacted]

[redacted]

[redacted]

Principle	Meaning in application

[redacted]



Applied together, the two parts provide a consistent approach to reviewing proposed harm mitigations with attention to foundational values. Every mitigation action must advance one or more of the core platform governance values and be consistent with the balancing principles of necessary, proportional and legitimate.

[redacted]

[redacted]

Conclusion

The Platform Principles-based Framework provides the analytical spine through which Snap identifies, investigates and mitigates systemic risk across Snapchat's in-scope services. The Framework's combination of core platform governance values, balancing principles and risk-based metrics supports a consistent, evidence-informed and rights-respecting approach to platform governance, and directly supports the Article 35(1) proportionality analysis recorded in the Section 4 risk assessments and the Section 5 mitigation sub-sections.

[redacted]

[redacted]

6.2 DSA Compliance Function and Cross-Functional Working Groups

Introduction

Section 6.2 sets out the governance mechanisms supporting Snap's compliance with the Digital Services Act, with particular focus on the requirements set out in Articles 11, 12, 13, 41 and 43 DSA. The framework operates as the principal organisational architecture through which Snap's substantive risk-mitigation work - recorded in detail across Sections 4 and 5 of this Report - is overseen, monitored, and adjusted on a continuing basis.

The framework is structured around four principal roles:

[redacted]

2026 DSA governance leadership structure

[redacted]

Roles and Responsibilities

The four principal roles within Snap's DSA governance framework are summarised below. Each role's substantive responsibilities are described in detail in the sub-sections that follow.

[redacted]

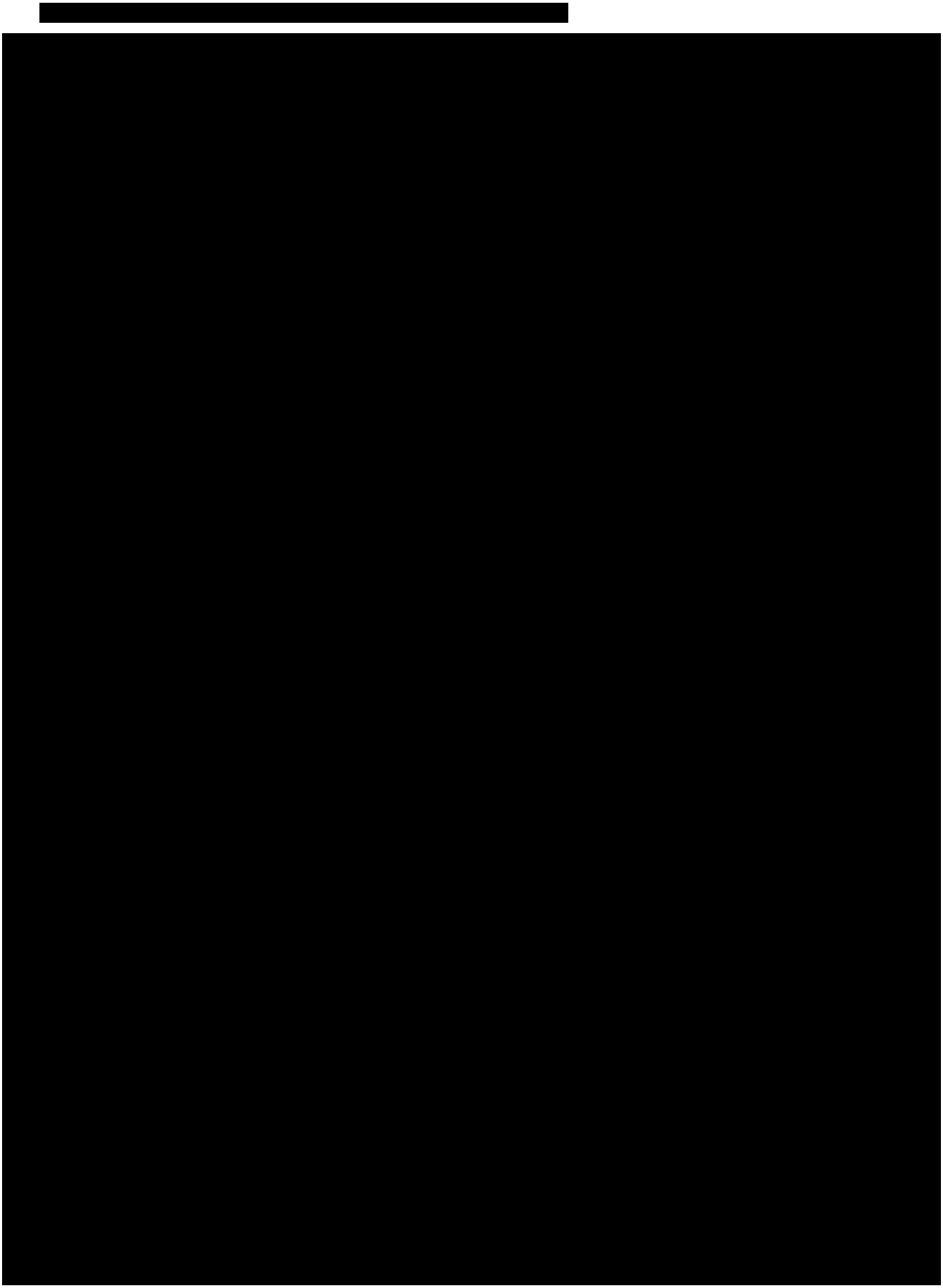
[redacted]

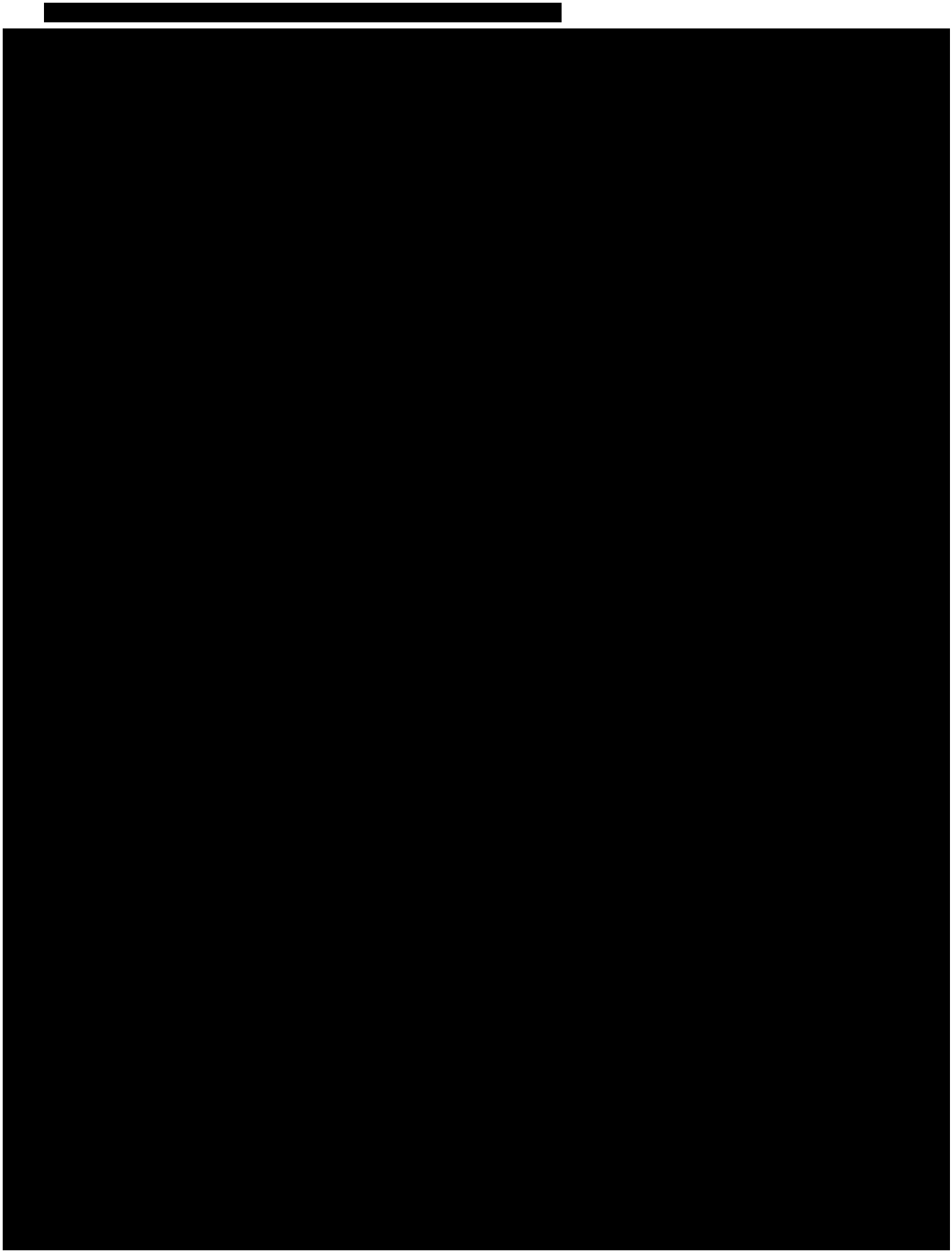
[redacted]

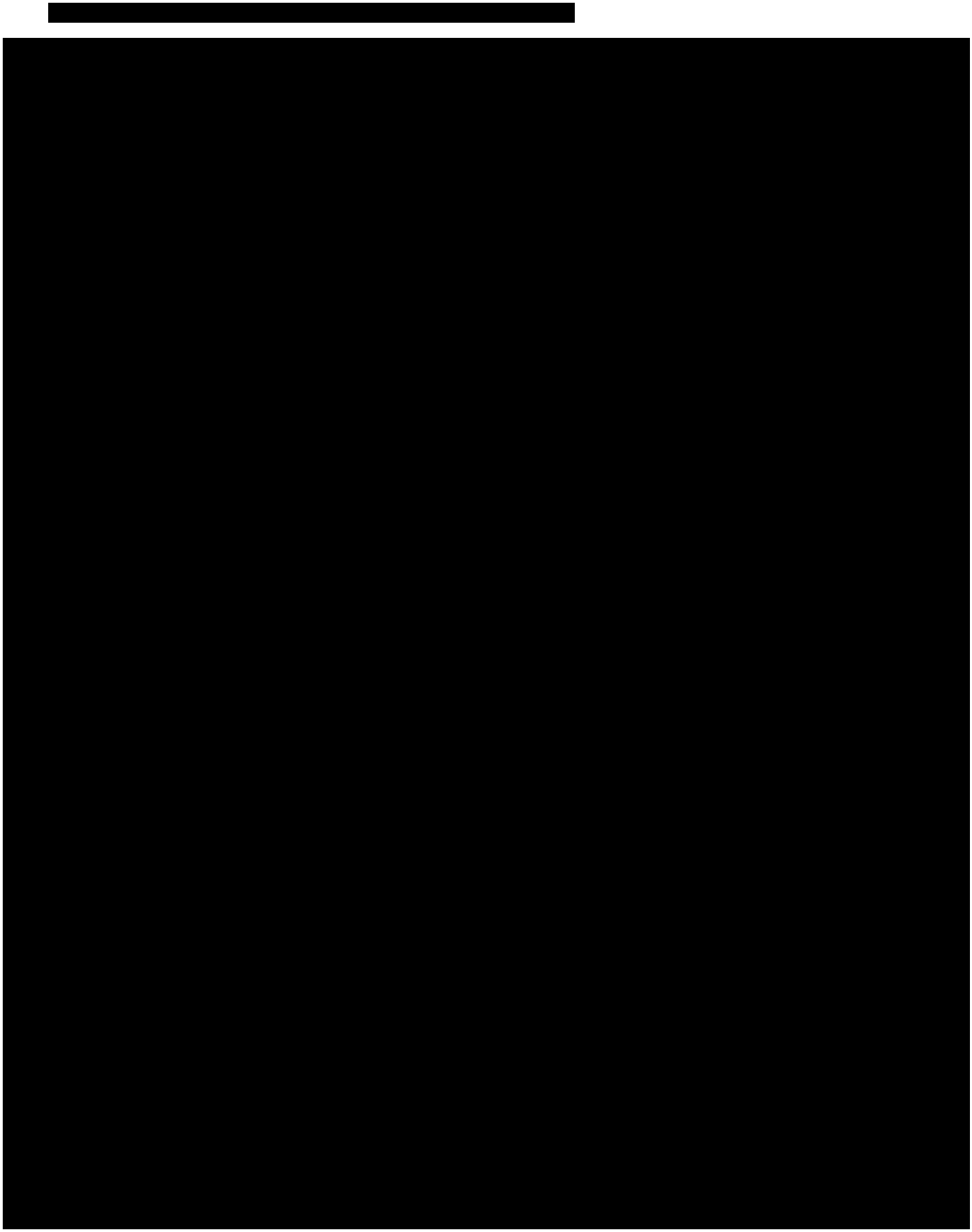
DSA Independent Compliance Function

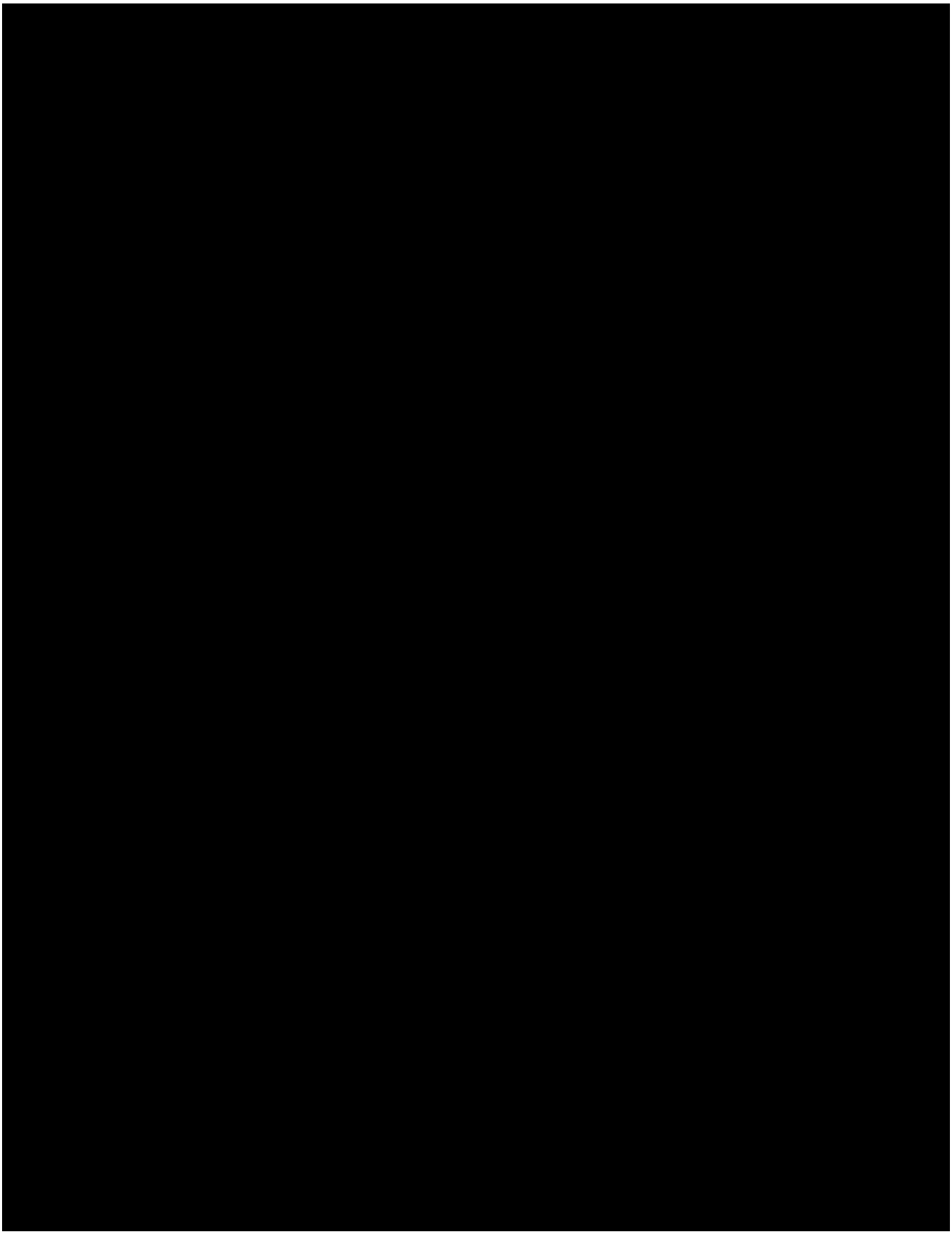
Snap has established an Independent Compliance Function (the "ICF") as part of the Snap Legal team, separate from Snap operations, with the authority, stature, and resources required for compliance with Article 41 DSA. The ICF is responsible for coordinating, overseeing, and implementing Snap's Privacy and Regulatory Programme. The ICF provides oversight to ensure the necessary internal processes, resources, testing, documentation, and supervision are in place for DSA compliance, and monitors Snap's compliance with the DSA, including the identification and mitigation of risks associated with Snap operations.

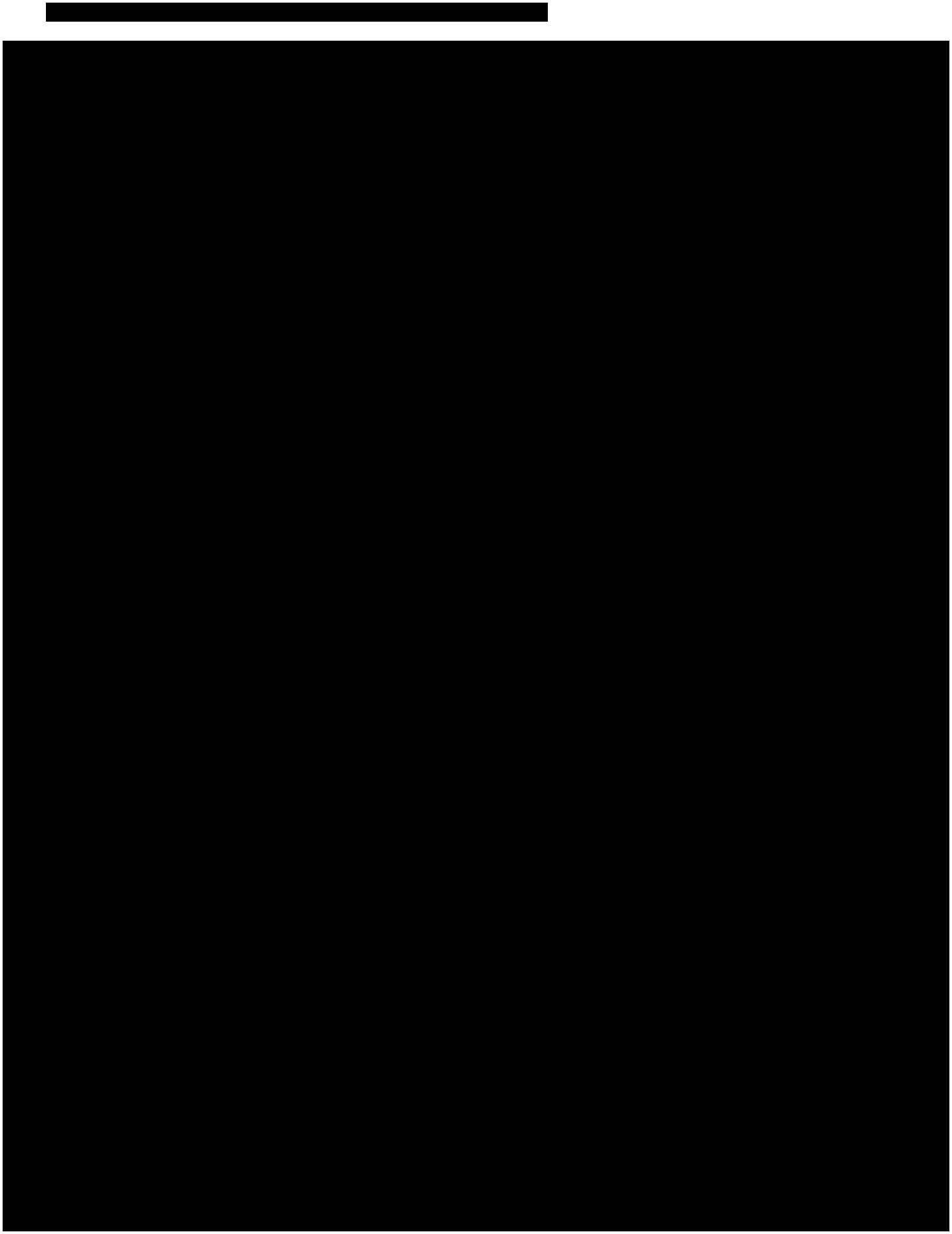
Mapping of Art. 41(3) DSA Compliance responsibilities Snap's compliance function extends beyond the formally designated roles within the Independent Compliance Function in operational terms. [redacted]

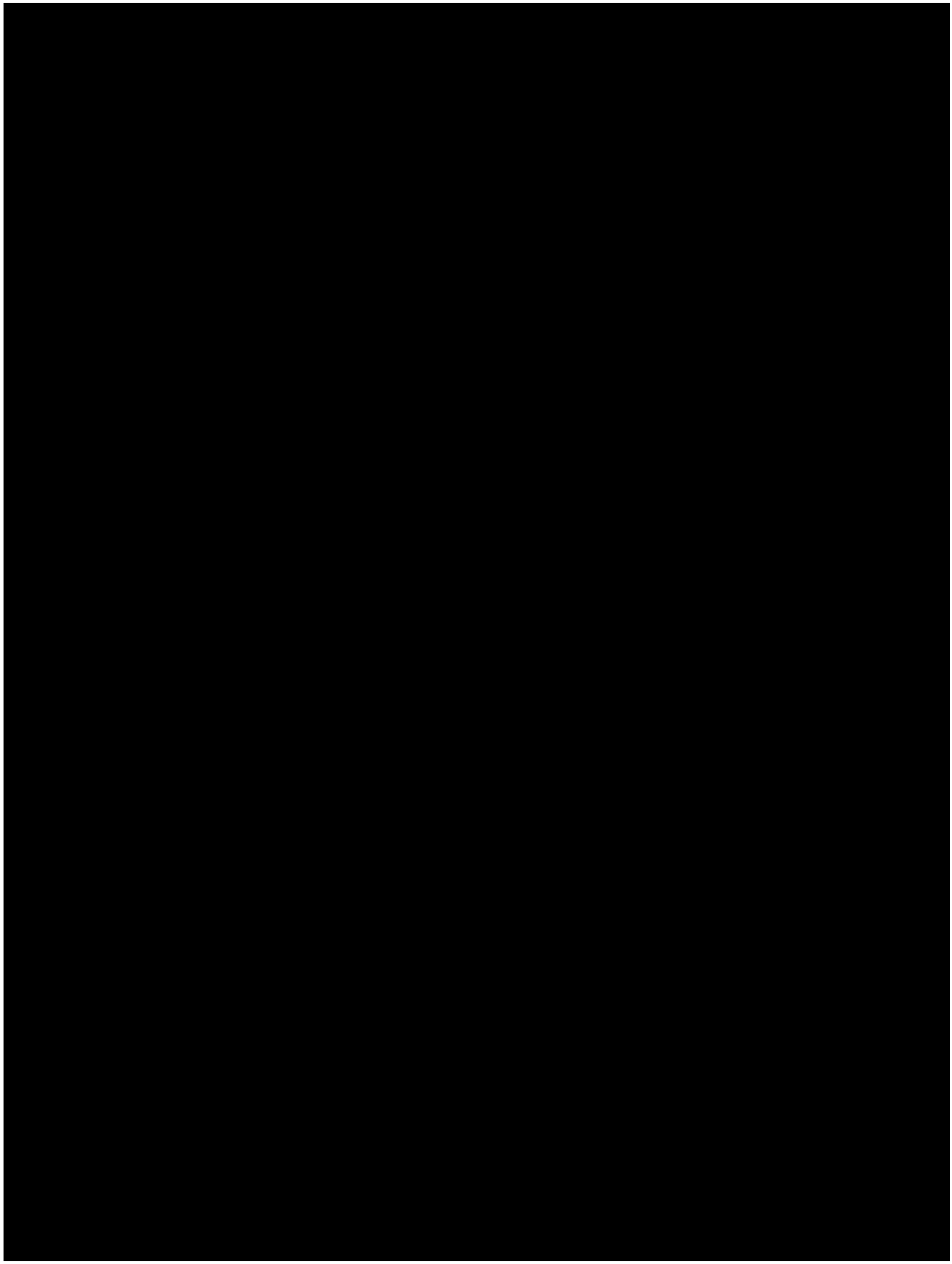












[redacted]

[redacted]

Points of Contact (Articles 11, 12, 13 DSA)

Designation, publication, and change management

Upon designation, Snap's DSA Cross-Functional Governance Team established a process for designating and updating, in a timely manner, the publication of points of contact for regulatory authorities (Article 11 DSA), recipients of the service (Article 12 DSA), and the Legal Representative (Article 13 DSA), in a publicly accessible location. The DSA Cross-Functional Governance Team reviews points of contact and the Legal Representative designation on an annual basis or when contact information changes; updates are coordinated with the Snap web team for prompt publication.

Points of contact and the Legal Representative designation are made publicly available in all official languages of the EU and are easily accessible at: **https://values.snap.com/privacy/transparency/european-union**. The mapping below records the substantive operational implementation of each point of contact:

[redacted]

[redacted]

[redacted]

DSA Supervisory Fee (Article 43 DSA)

Article 43 DSA establishes a supervisory fee payable by VLOPs and VLOSEs to support the European Commission's costs of supervision and enforcement under the Regulation. Where a supervisory fee is payable, Snap will pay it in accordance with the applicable Commission decision and payment instructions.

[redacted]

Conclusion

Snap's DSA compliance and governance framework operates through the integrated four-tier architecture recorded in this Section: 1) the DSA Cross-Functional Governance Team as the

[redacted]

principal cross-functional operational forum; 2) [redacted] compliance-oversight function under Article 41 DSA; 3) the DSA Management Body [redacted] as the senior oversight tier, and 4) the Risk Advisory Partners controls-testing function.

As recorded across Section 4, Snap has concluded that the governance framework recorded in this Section, in combination with the substantive mitigations recorded across Section 5 and the ongoing risk-detection processes recorded across Section 6, is reasonable, proportionate and effective for the systemic risks identified for Snapchat's in-scope services. The framework will continue to be subject to ongoing review and refinement supporting continuous calibration to the operational and regulatory environment.

6.3 Privacy and Safety by Design

Introduction

Section 6.3 sets out Snap's Privacy and Safety by Design framework, the principal cross-functional process through which Snap assesses privacy, safety, and DSA systemic-risk implications of new products, features, and material changes prior to launch. The framework operates as the principal Article 34 DSA ongoing-risk-detection mechanism ("risks must be assessed prior to deploying functionalities that are likely to have a critical impact on those risks") and as the operational pathway through which the Section 4 risk-assessment outputs are kept current as the in-scope services evolve.

Privacy and Safety by Design is foundational to Snap's approach to designing and launching products. The framework was in place before the entry into force of the DSA and has since been integrated with Snap's broader DSA compliance programme, including the Digital and Data Impact Assessment ("DDIA") framework recorded below. The framework is reviewed by Snap's Cross-Functional Governance Team (recorded in Section 6.2) on an annual basis.

DSA Risk Management context

Snap recognises a duty to assess new products and functionalities prior to deployment to determine whether they are likely to have a critical impact on the risks identified in this Report and on the corresponding mitigations. The duty is required under Article 34 DSA, recommended in the European Commission's published Article 34 / 35 Guidance, and aligns with industry-standard risk management practices and adjacent legal frameworks including the GDPR (DPIAs and Legitimate Interest Assessments). The Privacy and Safety by Design review process operates as Snap's principal mechanism for satisfying this duty.

[redacted]

[redacted]

[redacted]

Privacy and Safety by Design review process

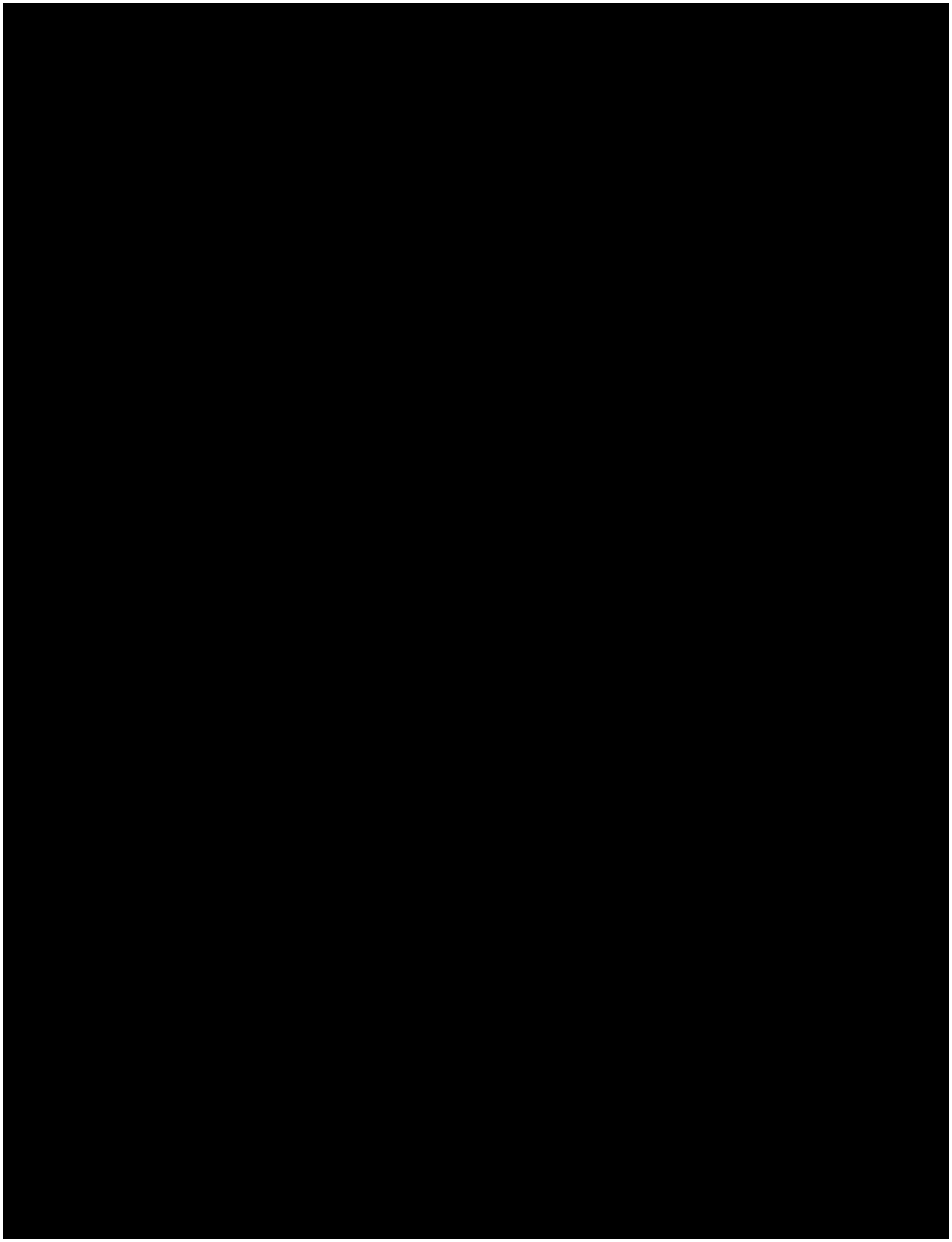
The Privacy and Safety by Design review process is a structured, cross-functional process embedded throughout the product development lifecycle. [redacted]

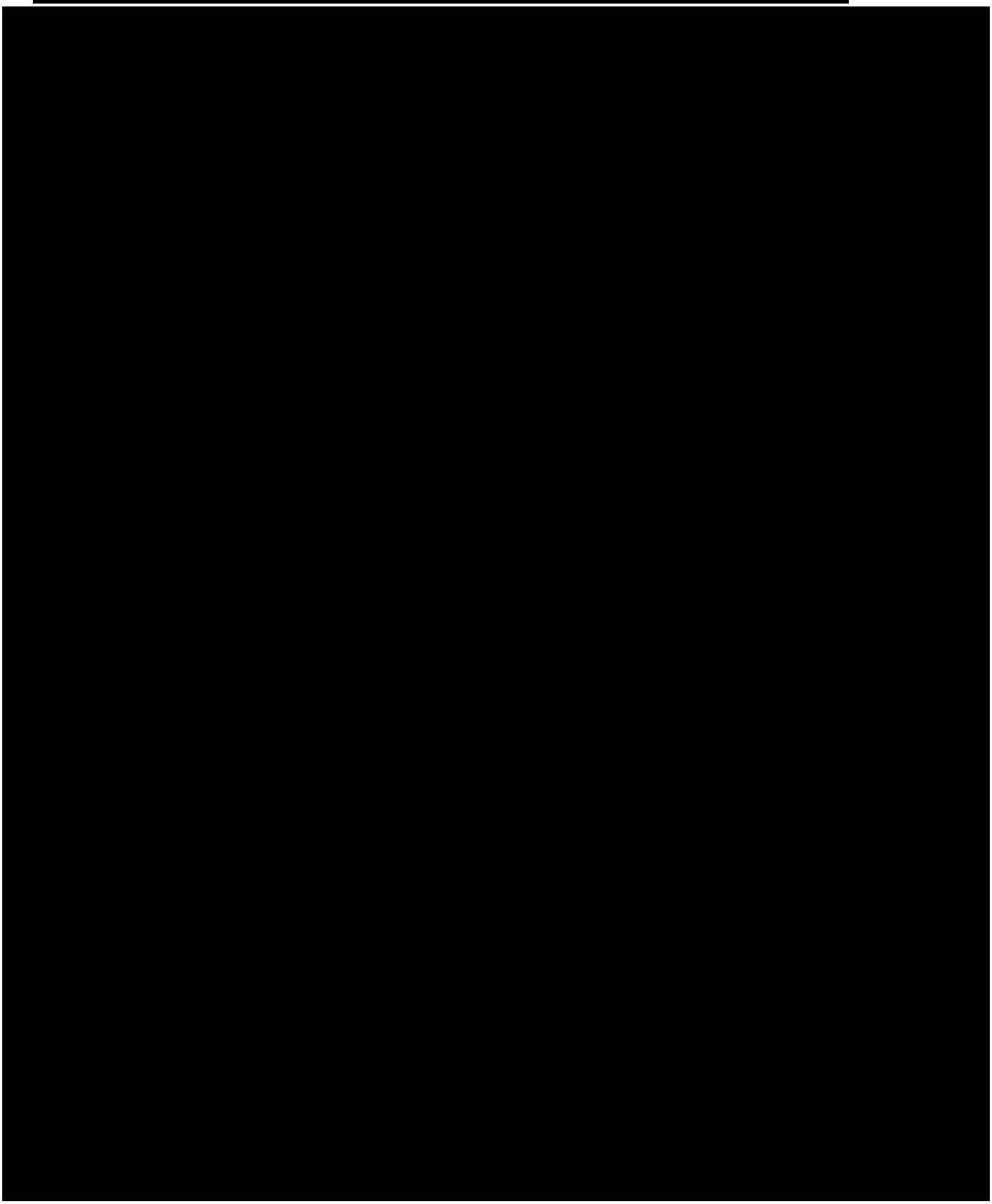
[redacted]

[redacted]

[redacted]

[redacted]





[redacted]

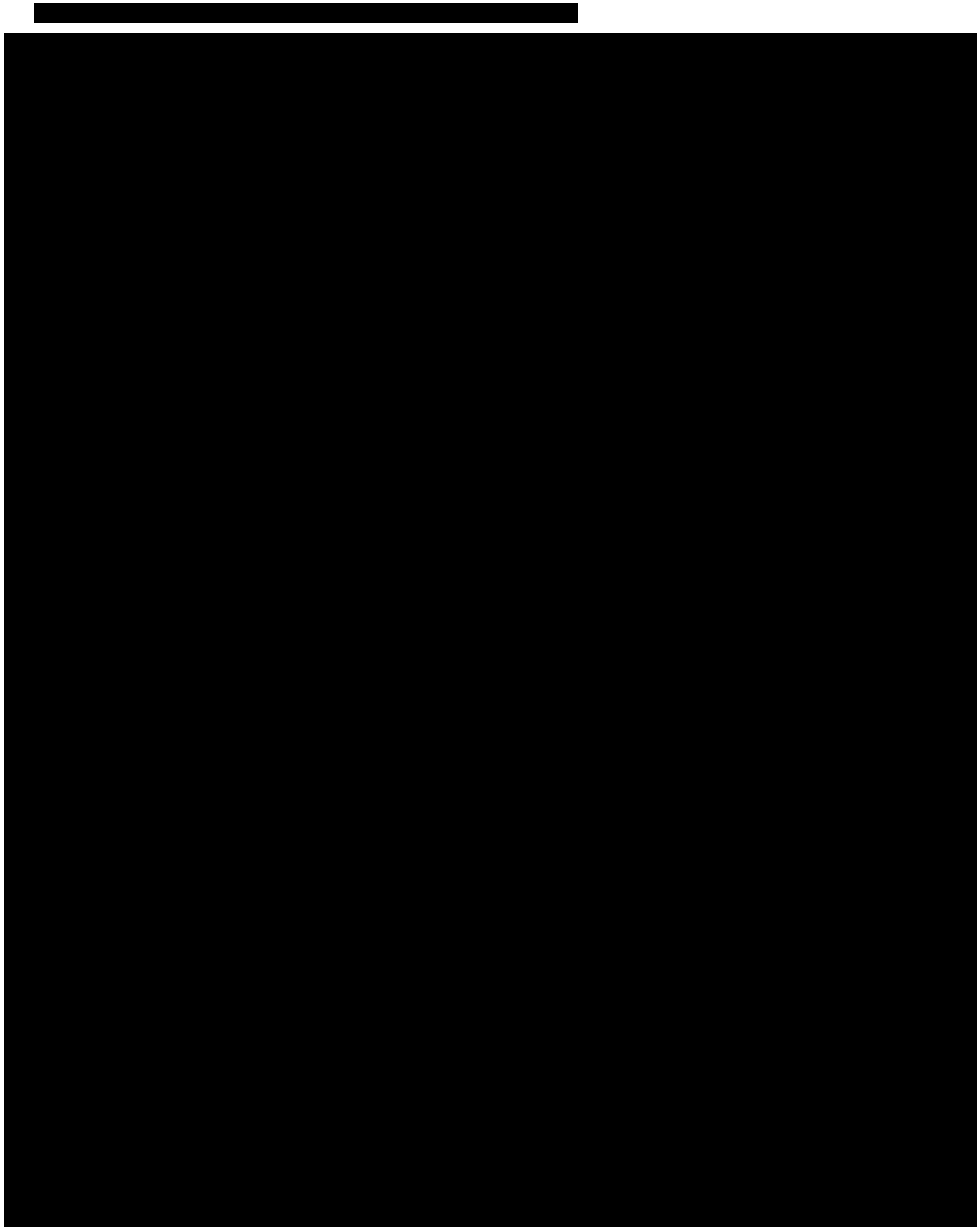
[redacted]

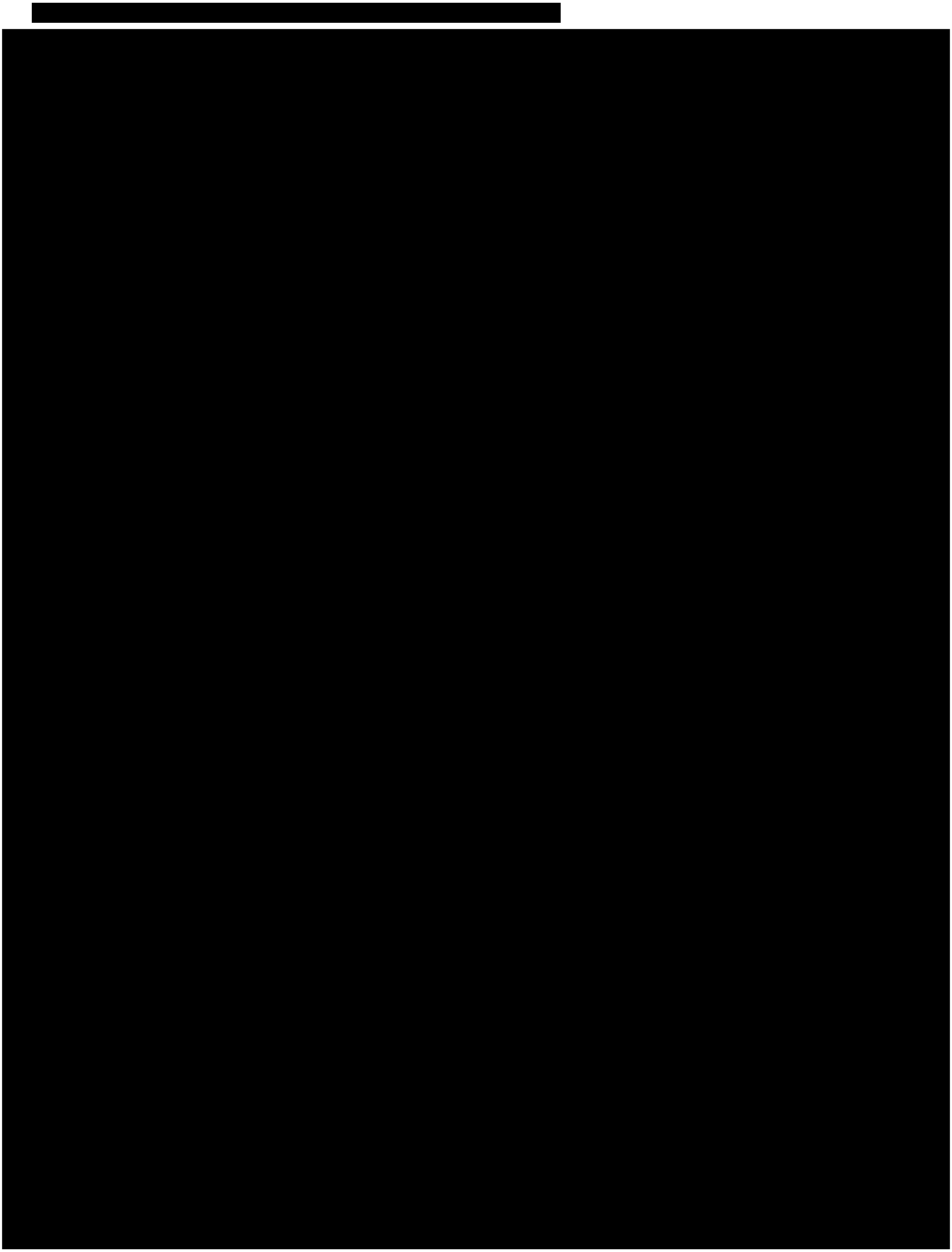
Digital and Data Impact Assessment (DDIA) Framework

Formal risk-assessment and mitigation obligations are an increasingly common feature of digital service regulation. In Europe, such obligations are imposed by the DSA; by the GDPR (in the form of Legitimate Interest Assessments and Data Protection Impact Assessments); and by the UK and several EU Member States through Age Appropriate Design Code (AADC) Assessments. To manage these overlapping European requirements alongside other global requirements in an efficient, effective, and operationalised manner, Snap operates an integrated Digital and Data Impact Assessment ("DDIA") framework that combines privacy, safety, and security obligations into a single risk-assessment instrument.

In each case, product reviewers use the template to record the aspects of the risk and mitigations particularly relevant to the feature to which the DDIA relates. The DDIAs in their entirety inform the design of Snap's recommender and algorithmic systems, the Terms framework, the content moderation systems, the enforcement framework, the systems for selecting and presenting advertisements, and considerations of data-related practices. The analysis from each DDIA is combined into this Report to enable a cross-platform view of risk.

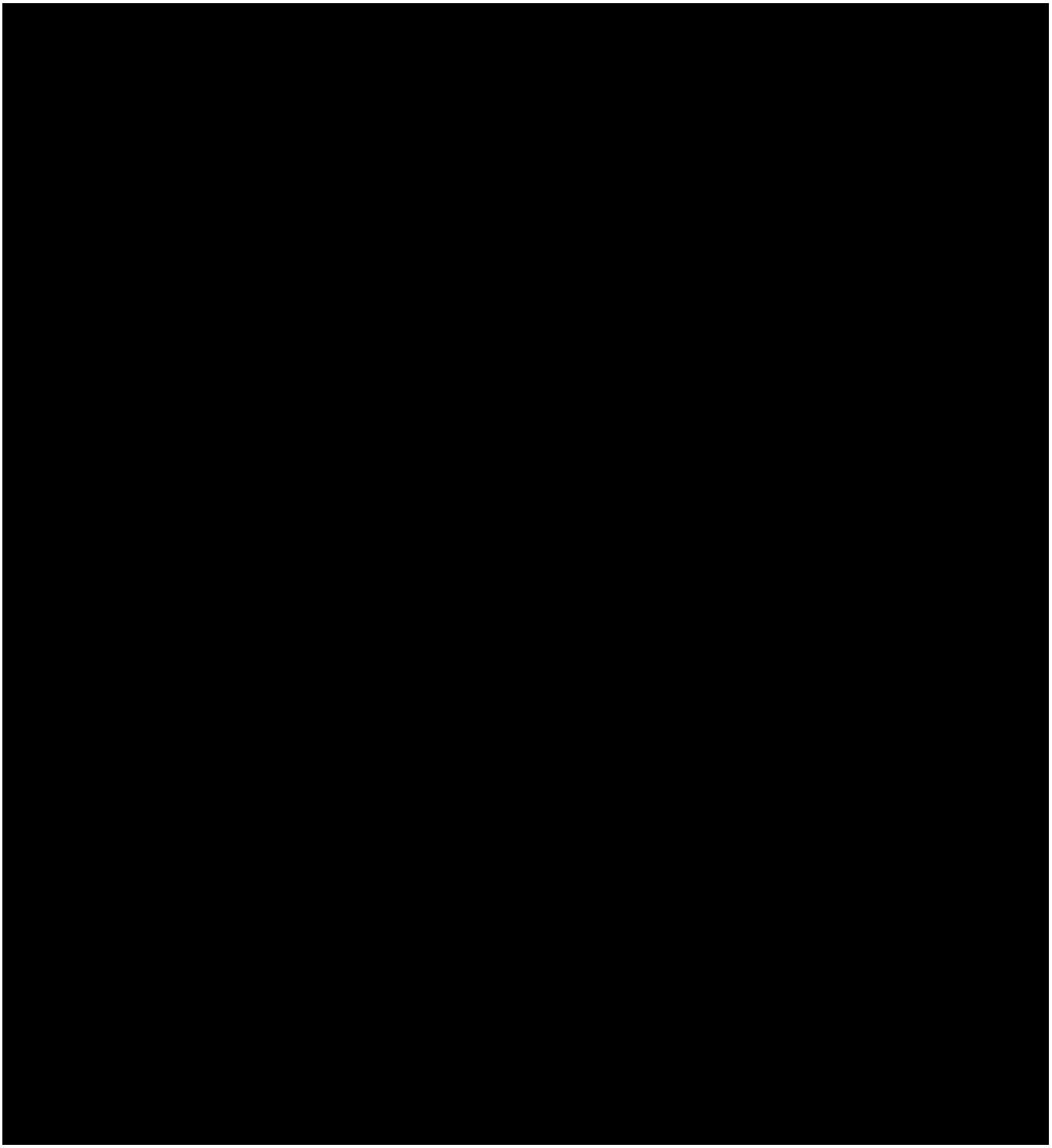
[redacted]





[redacted]

[redacted]



6.4 Prevalence Testing

Purpose and methodology

Prevalence testing is a key component of Snap's ongoing detection and management of illegal and other policy-violating public content. **Policy-Violating Prevalence (PVP)** estimates the proportion of views associated with content that violates Snap's Community Guidelines, using

[redacted]

random sampling of Snaps accessible through public surfaces. The results support identification of detection and enforcement gaps, prioritisation of mitigation work, and evaluation of the effectiveness of existing controls.

[redacted]

Updated PVP Methodology Going Forward

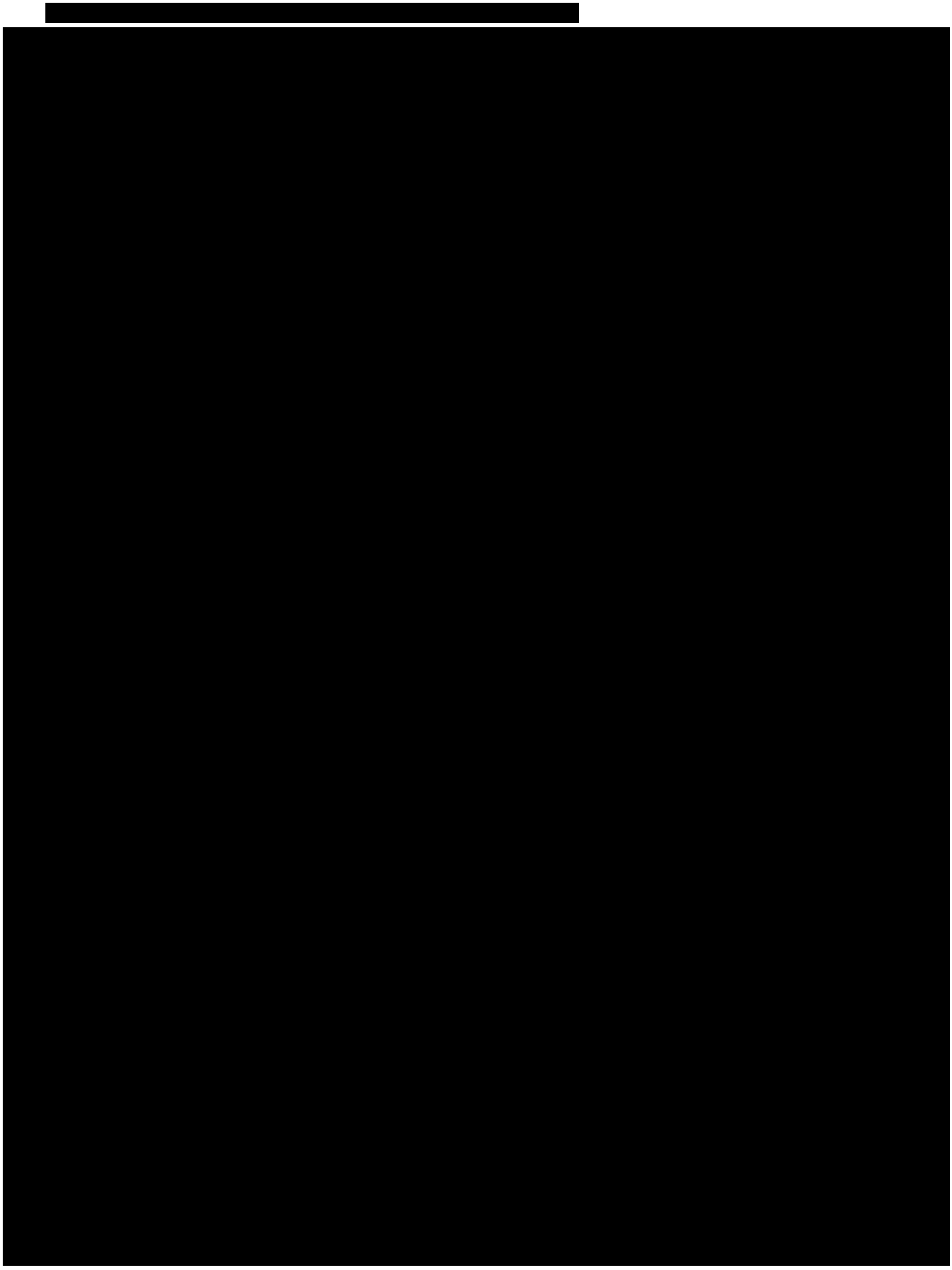
This year, Snap has refined its PVP calculation methodology and introduced a double-review process. Under this refined methodology a violation is counted only where it is confirmed through the applicable double-review process, meaning that unconfirmed violations are no longer calculated towards PVP. This removes sampled items that were initially identified as potentially violating but were not confirmed on subsequent review.

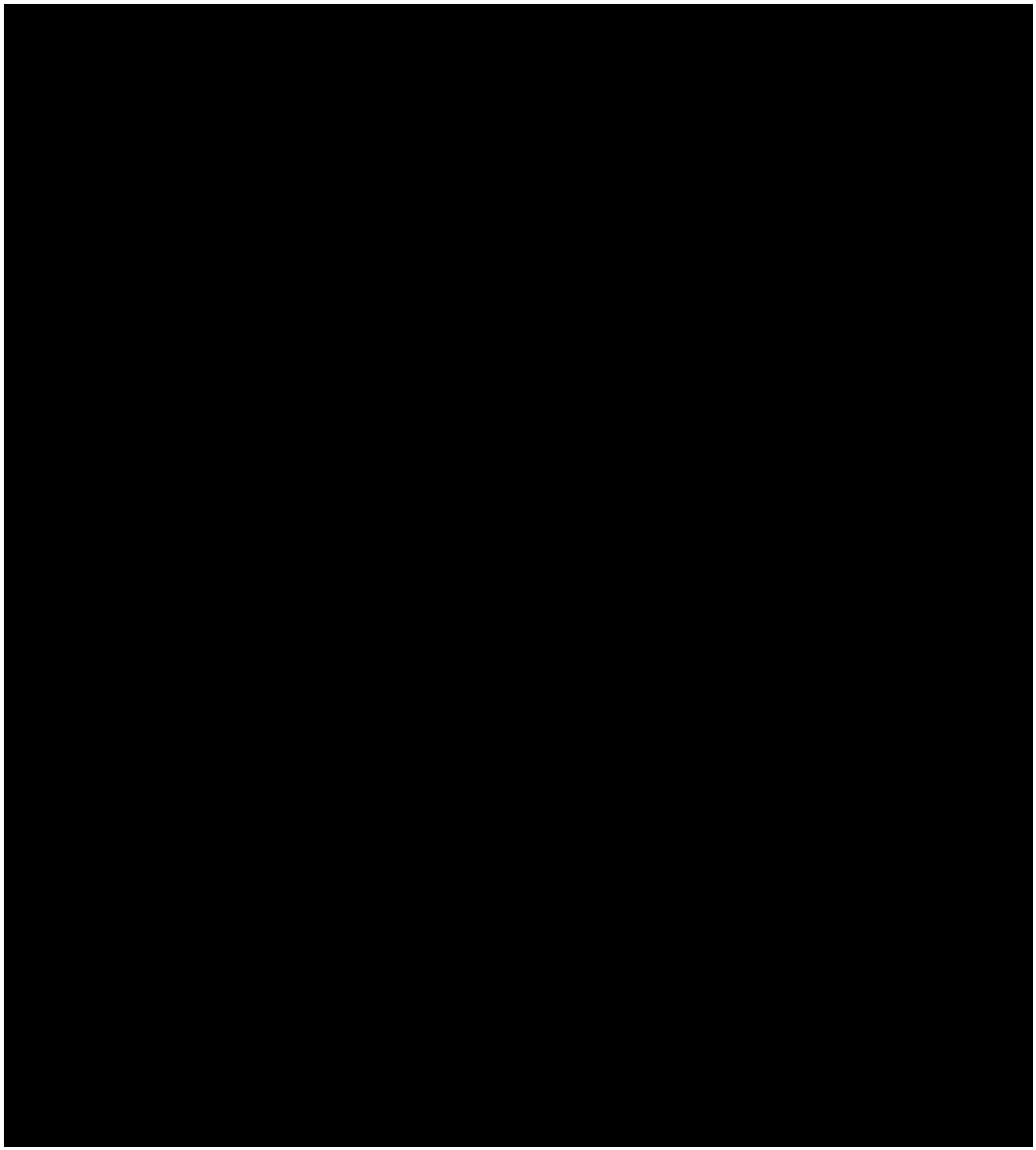
To preserve comparability with prior assessment periods, the 2026 PVP figures used in this Report are calculated using the overinclusive methodology to allow for a fair year-on-year trend analysis in this Report. In addition to including the old methodology, Snap has also included the 2026 PVP rates using the new methodology.

Going forward, Snap will use the refined double review methodology, and the 2027 Report PVP metrics will be compared against the updated 2026 metrics.

Headline results

[redacted]





Violative View Rate by audience

The EU dashboard also reports a six-month-average Violative View Rate (VVR) by enforcement reason for all users and separately for minor viewers. VVR provides a complementary exposure measure

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Operational use and conclusion

Prevalence-testing outputs are used to identify abuse patterns and control gaps, improve proactive detection models and rules, refine moderator guidance and training, calibrate enforcement, and prioritise product and operational mitigation work. [redacted]

6.5 External Request Monitoring and Review

Introduction

Section 6.5 sets out Snap's framework for monitoring and reviewing external requests received across the operational surface of Snapchat's in-scope services. The framework operates as one of the principal Article 35(1)(f) DSA ongoing-risk-detection instruments, supporting the substantive Section 4 risk-assessment determinations and the Section 5 mitigation framework through

[redacted]

structured engagement with the data streams generated by user-facing reporting, regulatory engagement, and law-enforcement coordination.

The framework integrates two principal categories of external request data:

- the data published in Snap's semi-annual Transparency Reports (Article 15 DSA, with the substantive framework recorded in Section 5.3); and
- additional internally-monitored external request streams that support the substantive risk-assessment cycle.

Both categories operate together to inform the analytical determinations recorded across Section 4, the operational refinements recorded across Section 5 and Section 6.3, and the continued evolution of the substantive risk-mitigation framework.

Transparency Report data streams

Snap produces a semi-annual Transparency Report (issued every six months), recording substantive enforcement and operational data across three principal data streams. The substantive framework for Snap's Article 15 / 24 DSA transparency reporting is recorded in Section 5.3. The data streams reviewed under this Section 6.5 framework are:

Transparency Report data stream	Description and substantive cross-reference
Community Guidelines enforcement data	Substantive enforcement-action volumes by category, recorded across the twelve Category 1 substantive harm categories of Section 4.1 (CSAM, Hate Speech, Sale of Prohibited Products, Terrorist Content, IP Infringement, Adult Sexual Content, Harassment and Bullying, Self-Harm and Suicide, Violent or Dangerous Behaviour, Harmful False Information, Fraud and Spam, and Other Illegal Activities). The data informs the substantive risk-assessment determinations recorded in the relevant Section 4 sub-sections and supports the Section 5.4 (Content Moderation) and Section 5.5 (Enforcement) frameworks.
Law-enforcement operations data	Article 9 DSA takedown order volumes; Article 10 DSA information request volumes; Article 18 DSA notification of criminal offences. Substantive operational detail recorded in Section 5.5 (Enforcement) under the Articles 9, 10 and 18 DSA framework, including the [redacted] and the International queue routing for fourteen named Member States. During the 2026 assessment period, Snap's Safety Operations Outreach function conducted direct engagement [redacted]

[redacted]

	[redacted]
Copyright and trademark data	Article 16 DSA notice and Article 17 DSA statement-of-reasons volumes engaging IP infringement, recorded substantively in Section 4.1.5 (IP Infringement). The H2 2024 baseline (1,296 copyright notices / 96.8% removal rate) is recorded in Section 4.1.5; 2025–2026 refresh figures are tracked through this monitoring framework.

Snap's Transparency Reports support insight into the substantive content moderation work, the law-enforcement engagement framework, and the IP-engagement work, supporting both external transparency (Article 15 / 24 DSA) and internal calibration (Article 35(1)(f) DSA).

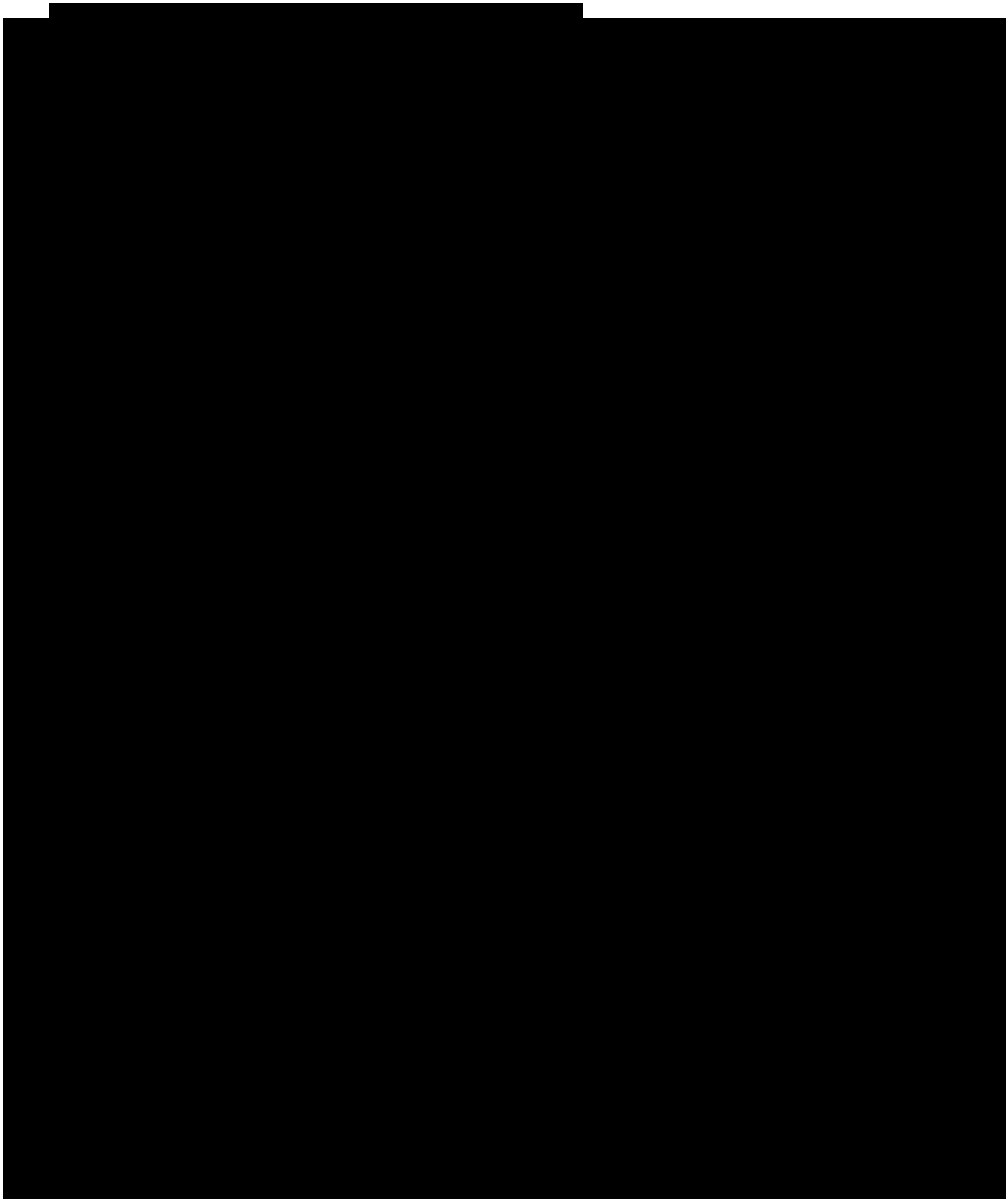
[redacted]

[redacted]

Additional external request streams

In addition to the Transparency Report data streams, Snap monitors additional external request streams that provide a continuous-monitoring framework for risk patterns. The principal additional streams are:

External request stream	Description and substantive cross-reference
Advertising review rejections	Pre-publication advertising rejection volumes, recorded substantively in Section 4.1.11 (Fraud and Spam) and Section 5.7 (Advertising Systems). The H2 2024 baseline (4,388,650 ads rejected; up from 1,825,144 in the comparable prior period) is recorded in Section 4.1.11; 2025–2026 refresh figures are tracked through this monitoring framework. Advertising review rejection volumes provide a leading-indicator metric for fraud and consumer-protection risk patterns (Section 4.2.7).
Advertising reporting and post-publication enforcement	Post-publication advertising reporting and enforcement volumes recorded substantively in Section 4.1.11 and Section 5.7. The H2 2024 EU baseline (23 reports confirmed and enforced post-publication) is recorded in Section 4.1.11.
Privacy, data-protection and DSA requests	Volumes of user-facing privacy and data-protection requests, recorded substantively in Section 4.2.3 (Right to Private Life) and Section 4.2.4 (Right to Data Protection). The H2 2024 baselines (1,617 privacy operations requests; 62 DPO queries) are recorded in those Sections; 2025–2026 refresh figures are tracked through this monitoring framework.
dsa-enquiries@snapchat.com queries	General DSA compliance queries received through Snap's dedicated Articles 11 and 12 contact point, including non-case-specific correspondence with certified ODS bodies. New Article 21 ODS disputes are submitted through the dedicated ODS webform and are recorded separately in the Article 21 ODS dispute stream below. Further detail on the general contact-point framework is recorded in Section 6.2.
General community support requests	Volumes of user-initiated community support requests across substantive categories, providing a baseline indicator of user-facing query patterns relevant to risk-detection across the Section 4 risk surface.
Article 21 ODS dispute stream	Disputes submitted by certified Article 21 DSA out-of-court dispute settlement bodies through Snap's dedicated ODS webform. Since October 2025, new submissions are ingested into Snap's



[redacted]

As recorded in the 2025 Report, Snap continues to review and use the external request data to support the conclusions reached in the substantive Risk Assessment work. The framework's outputs feed the Section 4 risk-assessment determinations and the Section 5 mitigation framework as part of a continuing calibration cycle, with the operational architecture supported by the Cross-Functional Governance Team's monthly meetings (Section 6.2), the Safety XFN's quarterly alignment cycle (Section 6.3), and the DDIA-update process (Section 6.3) for product or feature changes that engage the substantive risk surface.

Conclusion

Snap's external request monitoring framework operates as one of the principal Article 35(1)(f) DSA ongoing-risk-detection instruments. The framework integrates the data streams published through Snap's semi-annual Transparency Reports (Community Guidelines enforcement data, law-enforcement operations data, copyright and trademark data) with additional internally monitored streams (advertising review rejections; advertising reporting and post-publication enforcement; privacy, data-protection and DSA requests; the dsa-enquiries@snapchat.com query stream; general community support requests; and Article 21 ODS disputes). External request data is integrated into the substantive risk-assessment cycle through four principal applications: risk-assessment determination, risk-prioritisation calibration, monitoring, and continuous improvement of the mitigation framework.

As recorded across Section 4, Snap has concluded that the external request monitoring framework recorded in this Section, in combination with the broader Section 6 ongoing-risk-detection architecture and the substantive Section 5 mitigations, is reasonable, proportionate and effective for the systemic risks identified for Snapchat's in-scope services.

6.6 Digital Well-Being Index (DWBI) Initiative

Introduction

Section 6.6 sets out the Digital Well-Being Index ("DWBI") initiative, Snap's continuing cross-platform research programme on Generation Z teen and young adult online experience. The DWBI operates as one component of Snap's Article 35(1)(f) DSA ongoing-risk-detection and internal-calibration framework in respect of the substantive minor-protection and well-being

framework, supporting the analytical determinations recorded across Section 4.2.6 (Children's Rights), Section 4.4.3 (Negative Effects on Minors), and Section 4.4.4 (Negative Effects on Physical and Mental Well-Being).

The DWBI is a platform-agnostic research instrument: it investigates the digital risks and potential harms that teens and young adults self-report as encountering across all platforms, services, and devices, not just Snapchat. [redacted]

The research has continued on an annual basis. The 2024 research cycle–Year Three of the DWBI–was published in February 2025 and was a subject of the March 2025 expert event described below. The 2025 research cycle, Year Four, was published in February 2026. The six-country DWBI increased by one point from 63 to 64. Forty-five percent of respondents were classified as "Thriving" online and 13% as "Flourishing," while 36% were classified as "Middling" and 6% as "Struggling" – the strongest results across these categories thus far.

The DWBI is published through Snap's transparency framework. [redacted]

Research design and methodology

The DWBI research design is structured to support cross-platform analytical engagement with the teen and young adult online-experience risk surface. The principal design elements are:

Element	Description
Research scope	Cross-platform research on the digital risks and potential harms that teens and young adults are reportedly encountering across all platforms, services and devices, not limited to Snapchat. The cross-platform scope is structurally important: it positions DWBI findings as evidence on the broader online environment in which Snapchat operates, supporting the substantive Section 4 risk-assessment work without conflating Snapchat-specific data with industry-wide patterns.

Geographic coverage	Six countries: Australia, France, Germany, India, the United Kingdom, and the United States. The selection includes three of the largest European countries (France, Germany, the United Kingdom), two of which (France, Germany) are in the European Union. The six-country selection supports analytical cross-comparison while preserving sample volumes adequate for substantive interpretation.
Population coverage	Generation Z and Alpha teens (aged 13-17) and young adults (aged 18-24), supplemented by parents of teenagers aged 13 to 19. The three--cohort design supports analytical engagement with both the youth-self-report and parent-perception dimensions of the substantive risk surface.
Research cadence	Annual research cycle. The inaugural index was published on Safer Internet Day 2023, drawing on Spring 2022 research. The 2024 research cycle–Year Three–was published in February 2025 and was the subject of the March 2025 expert event described below. The 2025 research cycle–Year Four–was published in February 2026, with the six-country DWBI increasing from 63 to 64.
Public availability	DWBI findings are made publicly available through Snap's website (https://values.snap.com/), supplemented by per-country infographics for each of the six countries surveyed, an explainer, and the full research results. The substantive transparency framework supports both internal calibration and external academic and civil-society engagement.

Use of DWBI findings in risk-assessment cycle

DWBI findings help inform the substantive Section 4 risk-assessment cycle through structured cross-references to the relevant minor-protection and well-being considerations. The principal integration points are:

Substantive Section	DWBI integration
Section 4.2.6 (Children's Rights)	DWBI findings on teen and young adult online experience help inform the participation-and-expression analysis recorded in Section 4.2.6. The cross-platform scope of DWBI findings supports the analytical framing that excessively restrictive minor-targeted measures themselves can limit children's rights to expression, information and participation.

[redacted]

Section 4.4.3 (Negative Effects on Minors)	DWBI findings on minor users' encounters with online risk – across all platforms, help inform the substantive minor-protection analysis recorded in Section 4.4.3. The substantive operational refinements are calibrated against the broader teen-population research base that DWBI contributes to.
Section 4.4.4 (Negative Effects on Physical and Mental Well-Being)	DWBI findings on the well-being dimension of teen and young adult online experience are the most directly relevant Section 4 cross-reference. DWBI integrates with the broader academic and civil-society research base (LSE EU Kids Online; APA Health Advisory; Council of Europe Recommendation CM/Rec(2018)7; the Teen Research compendium of December 2025) to help support the substantive Section 4.4.4 determinations.
Section 4.1.6 (Adult Sexual Content / Sextortion)	DWBI findings on minor exposure to sexual content and on sextortion-pattern awareness inform the substantive sextortion sub-section analysis recorded in Section 4.1.6.
Section 4.1.1 (CSAM)	DWBI findings concerning reported exposure to sexual AI-generated imagery, including respondents' perceptions that some imagery may depict minors, inform the emerging-risk analysis concerning AI-generated CSAM identified at the March 2025 expert event, help inform the substantive CSAM analysis recorded in Section 4.1.1 and support the integration of generative-AI considerations into the ongoing risk-detection framework.

DWBI findings are integrated alongside the broader academic and civil-society research base recorded across the relevant Section 4 assessments. The cross-platform scope of DWBI findings provides analytically discrete evidence on the online environment in which Snapchat operates, supporting the substantive risk-assessment determinations without conflating Snapchat-specific operational data with industry-wide patterns.

[redacted]

[redacted]

[redacted]



Conclusion

The Digital Well-Being Index initiative operates as one component of Snap's Article 35(1)(f) DSA ongoing-risk-detection and internal-calibration framework in respect of the substantive minor-protection and well-being framework. The cross-platform study design supports analytically discrete evidence on the broader online environment in which Snapchat operates, informing the substantive Section 4 risk-assessment determinations, in particular, the analytical work in Section 4.2.6 (Children's Rights), Section 4.4.3 (Negative Effects on Minors), and Section 4.4.4 (Negative Effects on Physical and Mental Well-Being), and supporting the integration of generative-AI considerations into the framework as identified at the March 2025 expert event.

[redacted]

As recorded across Section 4, Snap has concluded that the DWBI initiative, in combination with the broader Section 6 ongoing-risk-detection architecture and the substantive Section 5 mitigations, supports the conclusion that Snap's mitigations for the systemic risks identified for Snapchat's in-scope services are reasonable, proportionate, and effective. The framework will continue to operate during the 2026 assessment period with expected continued annual research cycles, continued engagement with civil-society organisations focused on children's digital rights, and continued integration of substantive findings into the Section 4 risk-assessment cycle and the Section 5 mitigation framework.

6.7 Snap Advisory Groups

Introduction

Section 6.7 sets out the principal advisory and external expert engagement frameworks supporting Snap's substantive risk-detection work. The framework operates as one of Snap's principal Article 35(1)(f) DSA ongoing-risk-detection instruments through structured engagement with subject-matter experts, civil-society organisations, youth advocates, and EU/Member State stakeholder networks.

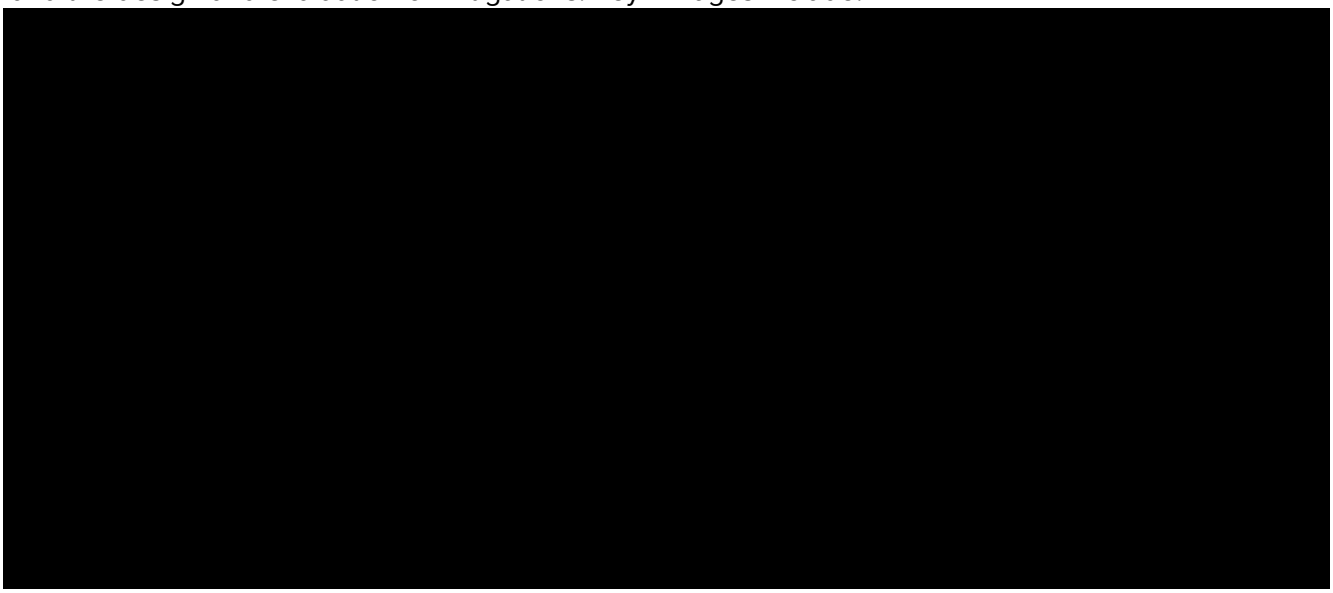
The framework is structured around three principal pillars:

[redacted]

Relationship with Snap's compliance program

The programs and external expert engagements recorded in this Section were established as part of Snap's broader safety, well-being, and stakeholder-engagement efforts and, where relevant, now also support Snap's DSA compliance and risk-governance architecture. They provide structured input from civil-society and academic experts on the systemic risk topics

addressed in this Report and help inform both the ongoing refinement of Snap's risk assessment and the design and evaluation of mitigations. Key linkages include:



The substantive risk analysis and mitigations are recorded in the relevant assessments in Section 4 and Section 5; this Section 6.7 records the advisory and external expert engagement framework that supports the broader risk-detection architecture.

Safety Advisory Board (SAB)

Snap launched the Safety Advisory Board (SAB) in 2018 and reconstituted and expanded it in 2022, with the aim of growing a body of expertise reflecting a diversity of geographies, safety-related disciplines, and areas of specialization. For this reconstitution, Snap initiated an application process inviting experts and individuals from around the world to formally express their interest in providing guidance and direction to Snap on "all things safety."

The SAB was initially established to educate, challenge, raise issues with, and advise Snap on how best to keep the Snapchat community safe, and to operate as a counterbalance to the online harms-dominated external landscape. Where appropriate, the SAB provides feedback on new products, features, policies, and initiatives before they are launched or released. The SAB and its individual members do not act as a representative or spokesperson for Snap, but rather as a collection of independent voices. The framework supports the substantive direction of Snap's approach to important safety issues and provides Snap with strategic safety-related advice and guidance as Snap grows.

SAB composition (2025–2026 cycle)

The current SAB composition is recorded below:

SAB element	Detail

[redacted]

[redacted]

SAB cadence and operating model

The SAB meets three times annually:

[redacted]

[165] https://values.snap.com/safety/safety-advisory-board.

[redacted]



In addition to this working-group exercise, the SAB received updates on broader Snap initiatives, including the youth-focused Council for Digital Well-Being programme recorded below. SAB input and feedback from the June 2025 meeting was memorialised and shared with Snap executives and others unable to attend in person.

Additional SAB meetings occurred in October 2025, February 2026, and July 2026, with discussion topics including Family Center, reporting flows, and age assurance.

Snap Council for Digital Well-Being (CDWB) Program

Snap formally launched the Council for Digital Well-Being (CDWB) in January 2024 as a pilot youth-focused engagement programme in the United States. The Council is designed to encourage safer online habits and practices among teens, with the objective of supporting young people in championing their knowledge and insights in their schools and communities. The CDWB operates as a complementary instrument to the SAB, providing a direct teen-cohort engagement channel alongside the broader expert advisory framework.

Inaugural US programme

For the first iteration of the programme, applications were opened to US-based teens aged 13 to 16. The inaugural cohort was selected and announced in May 2024. Following two virtual monthly meetings, the cohort travelled to Snap headquarters in Santa Monica, California, in July 2024 for the first Council Summit. The two-and-a-half-day Summit consisted of breakout sessions for the

teen cohort and their parents/chaperones; full-group discussions; guest speakers; and a 90-minute “speed-mentoring” session with 18 Snap employees representing different roles and teams.

The Summit yielded substantive engagement on topics including online pitfalls, parental tools, and the relationship between digital and in-person social dynamics. Outreach material to support continued ambassador work was provided to participants. The inaugural US programme was followed by a second US cohort cycle during the 2026 assessment period. In May 2026, Snap selected the 2026/2027 US Council for Digital Well-Being cohort. The cohort comprises 19 teens and is intended to provide continuing youth input on digital well-being and online-safety issues, including reporting of online harms, social-media restrictions, and the opportunities and risks associated with artificial intelligence. Shortly after the close of the assessment period, in July 2026, the cohort and their parents/chaperones attended an in-person Summit at Snap’s Santa Monica headquarters, that included workshops, breakout sessions, project work, guest-speaker sessions, and engagement with Snap employees.

European Council for Digital Well-Being (E-CDWB)

Following the inaugural US programme, Snap launched a European programme (the “E-CDWB”), with the selection process completed in 2025. The first European Teen Council for Digital Well-Being Summit was held in Amsterdam at the end of June 2025. The E-CDWB is composed of 14 teens from 10 countries[166].

The Amsterdam Summit hosted the teens and their parents/chaperones, with the substantive programme covering their online life experiences, the things they enjoy, the value they derive from online platforms, the worries they hold, and their ideas for the future. The E-CDWB provides an ongoing channel for youth input through regular discussions focused on online safety and digital citizenship, while also supporting members in developing leadership, advocacy, teamwork, and peer-mentoring skills. In addition to the summit, the program includes monthly calls and the year-plus-long program will conclude in August 2026 with a capstone event. We anticipate establishing a second European cohort in 2027. Feedback from the Council forms part of Snap’s broader youth-engagement and ongoing risk-detection work relevant to the minor-protection and digital-well-being frameworks recorded in Sections 4.4.3 and 4.4.4.

Australia programme

Snap also operated a youth-focused Council for Digital Well-Being programme in Australia on a timetable broadly mirroring the European programme. The Australian programme concluded in June 2026.

External engagement framework

[166] https://values.snap.com/safety/european-cdwb

[redacted]

Beyond the SAB and CDWB programmes, Snap engages with external experts and civil-society organisations on a continuing basis. The external engagement framework supports the substantive risk-detection work through several complementary channels.

Broader external expert engagement

Snap periodically consults a cadre of approximately 50 safety experts from around the world on significant new product features and functionality, policies, and initiatives. The substantive consultation framework supports calibration of the substantive risk-mitigation framework to the operational and academic environment. The framework operates alongside:

- internal trainings and learning sessions, with external experts invited to inform and educate Snap personnel working in safety disciplines on the broader risk landscape and Snap's potential exposure;
- ongoing collaboration with the US-based National Center for Missing and Exploited Children (NCMEC) on the topics of sextortion and CyberTip report quality (engaging the substantive frameworks recorded in Section 4.1.1 and Section 4.1.6); and
- smaller, executive-attended sessions with WeProtect Global Alliance, the Internet Watch Foundation (IWF), Thorn, and other cross-platform child-protection stakeholders (engaging the cooperation framework recorded in Section 5.12).

The substantive external partnership and relationship investment supports continuing internal knowledge and awareness of the overall risk landscape. The framework will continue to evolve during the 2026 assessment period.

July 2025 DSA Risk Assessment and Mitigation workshop

On 10 July 2025, Snap held its first dedicated DSA Risk Assessment and Mitigation workshop in Brussels and online. The session provided a detailed overview of Snap's 2024 DSA Risk and Mitigation Assessment Report and operated as a structured opportunity to collect feedback from safety NGOs based in Brussels and across EU Member States. The workshop served as a follow-up to the European Commission's workshop on systemic risk mitigation held on 7 May 2025, during which several stakeholders had expressed interest in more structured opportunities to provide input to platforms.

The workshop brought together 19 experts representing 14 NGOs. The participating organisations are recorded below, with cross-references to other engagement frameworks where applicable:

[redacted]

[redacted]

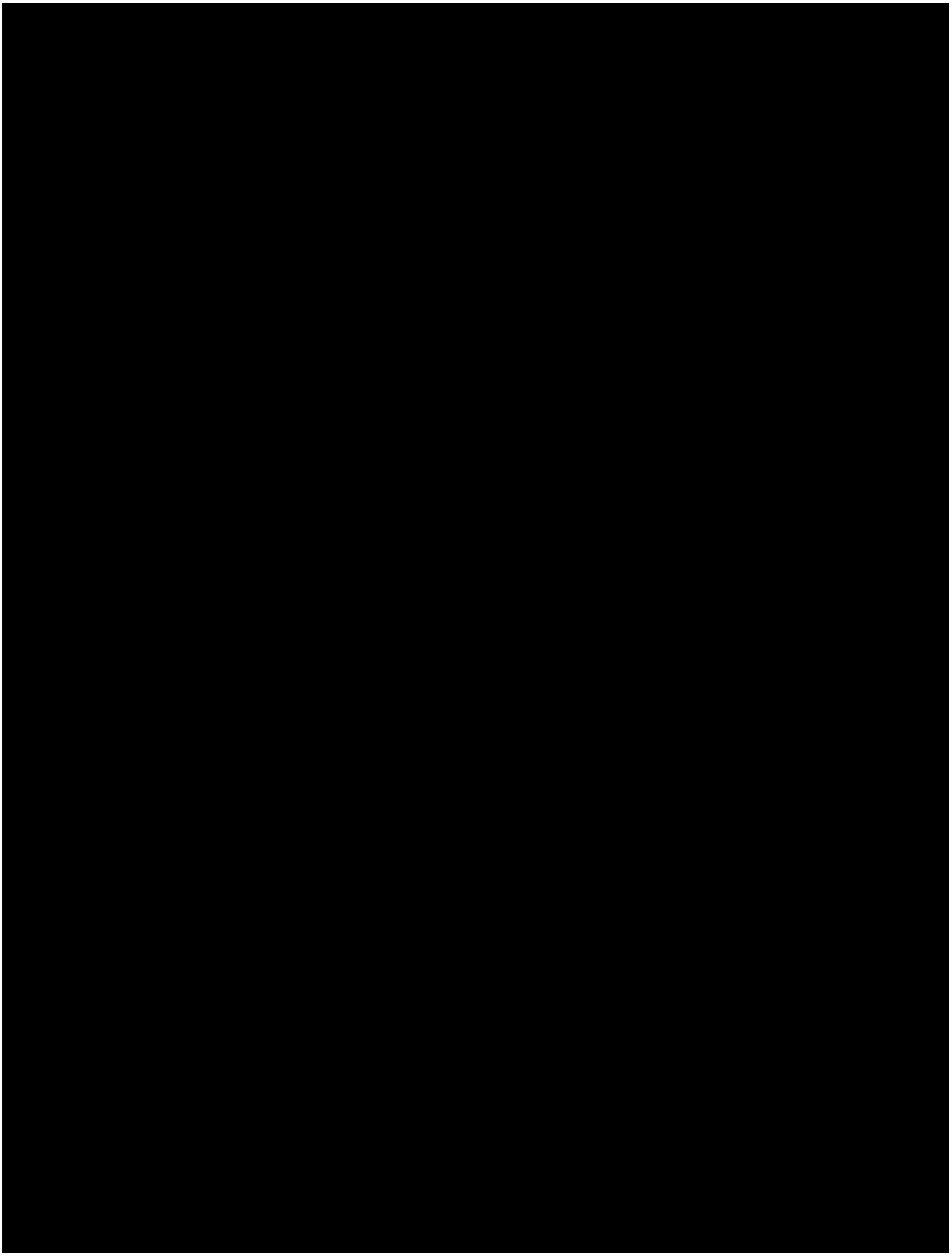
[redacted]

[redacted]

Feedback themes and 2026 Report integration

Participants engaged substantively with Snap's risk-assessment process and provided constructive feedback. The principal feedback themes, each integrated into the architecture of the 2026 Report, are recorded below:

Workshop feedback theme	Snap response in 2026 Report



[redacted]

[redacted]

September 2025 and March 2026 Expert Listening Roundtable Events

In September 2025, Snap hosted an expert listening roundtable event in London, and in March 2026 organised a similar event in Paris. In attendance were experts from different NGOs to discuss [redacted]

Engagement with External Researchers

Snap treats inquiries, research findings, and allegations from external researchers as important inputs into its systemic risk management processes. Where researchers identify potential risks or concerns, Snap engages directly with them where appropriate to better understand their methodology and findings, and relevant internal teams investigate the claims against Snap's own data, systems, policies, and product functionality. This process is intended both to identify substantiated issues requiring action and to assess whether external findings may be affected by methodological limitations, incomplete information, or incorrect assumptions.

Where an investigation identifies an actionable issue or opportunity for improvement, Snap uses the findings to inform appropriate product, policy, engineering, or operational measures. Conversely, where claims are not supported by Snap's investigation, Snap may engage with researchers to provide additional context, address methodological concerns, or correct inaccurate conclusions. This approach creates a feedback loop in which external research is actively assessed, validated, and, where appropriate, translated into risk mitigation measures and continuous improvements to Snap's services. [redacted]

Continuing engagement

Snap intends to continue regular engagement with participating NGOs to support transparency on how stakeholder input is reflected in ongoing risk-assessment work. Updates on future sessions and outcomes will be shared with the European Commission as the work progresses.

[redacted]

No follow-up DSA Risk Assessment workshop had been held as of the 2026 assessment cutoff following the 10 July 2025 Brussels and online session described above.

Conclusion

Snap's advisory and external expert engagement framework operates as one of the principal Article 35(1)(f) DSA ongoing-risk-detection instruments supporting the substantive risk-mitigation work recorded across this Report. The framework integrates structured engagement through the Safety Advisory Board (18 members from 11 countries, meeting three times annually), the Council for Digital Well-Being programmes (US, Europe, Australia), the dedicated July 2025 DSA Risk Assessment workshop with 14 EU/Brussels-based NGOs, and the broader external expert and civil-society engagement framework.

As recorded across Section 4, Snap has concluded that the framework recorded in this Section, in combination with the broader Section 6 ongoing-risk-detection architecture and the substantive Section 5 mitigations, is reasonable, proportionate, and effective for the systemic risks identified for Snapchat's in-scope services. The framework will continue to be subject to continuing operation and refinement during the 2026 assessment period, with continued SAB cycles, continued Council for Digital Well-Being programme operation across geographies, and continuing external expert and civil-society engagement supporting the substantive Section 4 / Section 5 frameworks.

6.8 Audit

Introduction

Section 6.8 sets out Snap's framework for compliance with the independent audit obligations under Article 37 DSA. The Article 37 DSA independent audit serves as the principal external assurance mechanism for assessing Snap's compliance with its obligations under Chapter III of the DSA. Audit findings, recommendations, and the corresponding audit implementation report are integrated into Snap's broader DSA compliance framework and inform continuous improvement across Snap's governance, risk assessment, controls, and mitigation activities.

Coordination with the Independent Compliance Function

The Article 37 independent audit is subject to substantive coordination and supervision by Snap's Independent Compliance Function, supported by the Risk Advisory Partners team, as described in Section 6.2. Consistent with Article 41(3)(c), Snap's DSA Compliance Officer performs the operational audit-supervision responsibilities of Snap's activities relating to the independent audit under the oversight of the Head of the Independent Compliance Function.

The substantive coordination framework includes:

[redacted]

The comments and recommendations identified through the 2025 audit were addressed in accordance with the audit implementation report.

2026 DSA audit

The 2026 DSA audit cycle, covering the period from 1 July 2025 through 30 June 2026, is ongoing as of the time when this report was prepared. [redacted]

Additional Independent Assurance

[redacted]

[redacted]

[redacted]

[redacted]

Conclusion

Snap completed the 2025 DSA audit cycle, including receipt of EY's independent audit report, preparation and submission of the corresponding audit implementation report, publication of the required public-facing materials, and implementation of the responses identified in the audit implementation report. The 2026 DSA audit cycle is currently in progress.

Snap considers the audit framework set out in this Section, together with the ongoing risk-detection and governance mechanisms described in Section 6 and the substantive mitigation measures described in Section 5, to contribute to a reasonable, proportionate, and effective framework for addressing the systemic risks identified in relation to Snapchat's in-scope services.

This framework will continue to operate throughout the 2026 assessment cycle. Outputs from the 2026 DSA audit will feed into Snap's broader DSA compliance programme and the ongoing governance, risk assessment, controls, and mitigation activities supported by the Independent Compliance Function and the DSA Cross-Functional Governance Team.

7. Conclusion

This Report has been prepared pursuant to Articles 34(1) and 35(1) DSA and is transmitted to the European Commission and the Dutch Digital Services Coordinator pursuant to Article 42(4) DSA. The Report sets out the results of: (i) the risk assessment conducted by Snap pursuant to Article 34(1); and (ii) the reasonable, proportionate and effective mitigation measures that Snap has put in place pursuant to Article 35(1) DSA, tailored to the specific systemic risks identified, and Snap's assessment of those measures.

The risk assessment conducted by Snap identified, analyzed and assessed in accordance with Article 34(1) DSA any systemic risks in the European Union stemming from the design, functioning or use of Snapchat's in-scope services. As required by Article 35(1) DSA, Snap has put in place mitigation measures that are reasonable, proportionate and effective, tailored to the specific systemic risks identified by its risk assessment, and has assessed those measures to confirm they

[redacted]

(continue to) meet that standard. The results of the risk assessment and mitigation review are set out in Section 4 of this Report. The specific mitigation measures put in place by Snap are further detailed in Sections 5 and 6 of this Report.

The Report demonstrates Snap has reasonable, proportionate and effective mitigation measures in place and continues to monitor a few areas, including:

- **Dissemination of illegal or violating content:** Since our previous Reports, Snap has observed a further reduction in the prevalence of content that is illegal or otherwise violating Snap's Terms being disseminated on Snapchat's in-scope services. Snap's overall Policy Violating Prevalence rate declined from [redacted] representing an approximately [redacted], consistent with continued improvements to detection mechanisms, reviewer training and enforcement processes across the public-content surfaces included in our prevalence testing. This year, all of Snap's illegal or other violating content categories were again observed from our testing to fall within our lowest relative likelihood category. Within this very low level of dissemination Snap:[167]
    - Continues to categorize two dissemination risk areas as falling within Level 1 risk prioritization for Snapchat's in-scope services: (i) child sexual abuse material, and (ii) sale of drugs. Snap has reasonable, proportionate and effective mitigation measures for both of these categories. As a result of these ongoing measures, these two categories are still assessed to fall within our lowest likelihood category of the risks identified by Snap.
    - Continues to categorize five dissemination risk areas as falling within Level 2 risk prioritization for Snapchat's in-scope services: (i) Sale of Weapons, (ii) Terrorist Content, (iii) Adult Sexual Content (sexual crimes), (iv) Harassment and Bullying and (v) Self-Harm and Suicide, due to the risk of serious harm each may cause. Snap has reasonable, proportionate and effective measures in place for all five of these categories. As a result of these measures all five are assessed to fall within our lowest likelihood category of the risks identified by Snap. There is no change in these determinations from our 2025 Report.
    - Continues to categorize eight dissemination risk areas as falling within Level 3 risk prioritization for Snapchat's in-scope services: (i) illegal Hate Speech, (ii) sale of prohibited products or services (excluding Drugs and Weapons), (iii) intellectual property infringements, (iv) other Adult Sexual Content, (v) Violent or Dangerous Behaviour, (vi) Harmful False Information, (vii) Fraud and Spam and (viii) content relating to Other Illegal Activities. Snap has reasonable, proportionate and effective measures in place for all eight of these categories. As a result of these measures all eight of these risk areas are assessed to fall within our lowest

[167] Risk prioritisation levels in this Report reflect the potential severity of the harm and the case-by-case prioritisation methodology described in Section 3, rather than a finding of elevated likelihood. A category may therefore be assigned a Level 1 or Level 2 prioritisation while falling within Snap's lowest relative likelihood category on the basis of observed prevalence. The same applies to the priority levels recorded for the fundamental rights, public security and public health categories below.

[redacted]

likelihood category of the risks identified by Snap. There is no change in these determinations from our 2025 Report.

- **Negative effects on EU Fundamental Rights:** Snap continues to categorize three risks to fundamental rights as falling within the highest priority level for Snapchat's in-scope services: (i) Human Dignity, (ii) Data Protection and (iii) Children's Rights. In the case of Data Protection, the highest-likelihood determination continues to be driven by the structural exposure inherent in the processing footprint rather than by the empirical incidence of complaints. Snap categorizes four risks as falling within the final priority level: (i) Right to Freedom of Expression and Assembly, (ii) Right to Private Life, (iii) Right to Non-Discrimination and Freedom of Religion and (iv) Right to Consumer Protection. This is a change from the 2025 Report, when the Right to Private Life was assessed to fall within the second highest priority level. Snap has reasonable, proportionate and effective measures in place for all of these categories. Snap continues to assess Snapchat's in-scope services against the Commission's Guidelines on Article 28 to ensure a high level of privacy, safety and security for minors and to assess whether further measures are needed to address risks to child rights.

- **Negative effects on Public Security:** Snap continues to categorize three risks to public security within the final priority level for Snapchat's in-scope services: (i) Negative Effects on Democratic and Electoral Processes, (ii) Negative Effects on Civic Discourse and (iii) Negative Effects on Public Security. Snap has reasonable, proportionate and effective measures in place for all of these categories.

- **Negative effects on Public Health:** Snap continues to categorize: (a) two risks to Public Health within the highest priority level for Snapchat's in-scope services: (i) Negative Effects on Minors; and (ii) serious negative consequences on Physical and Mental Well-Being; (b) one risk within the next highest priority level for Snapchat's in-scope services: Negative Effects on Gender-Based Violence; and (c) one risk within the final priority level for Snapchat's in-scope services: Negative Effects on Public Health. Snap has confirmed it has reasonable, proportionate and effective measures in place for all of these categories. As above, we continue to assess Snapchat's in-scope services against the Commission's Guidelines on Article 28 to ensure a high level of privacy, safety and security for minors and to assess whether further measures are needed to address risks to the protection of minors.

Snap remains focussed on further reducing the prevalence of illegal and otherwise violating content on Snapchat's in-scope services. [redacted]

[redacted]

Since our previous Reports, Snap has continued to evaluate its mitigation measures against new guidance from the Commission (in particular the implementation of the Commission's Guidelines on Article 28). During the 2026 assessment period, Snap introduced or strengthened a number of product safeguards, [redacted]. Snap has also updated the information on our measures to reflect the increased use of generative AI technology; shortly after the assessment period, in July 2026, Snap updated Spotlight's recommendation systems such that wholly AI-generated videos are no longer eligible for recommendation, which is noted in this Report as a subsequent development.

Since the previous Reports, Snap did not identify any deployed functionalities that were likely to have a critical impact on our assessment of risks and mitigations pursuant to Articles 34 and 35 of the DSA. In particular, the Friend-Only content profile for users aged 13–15, which began rolling out in June 2026, was assessed and not considered likely to have a critical impact on the risks identified pursuant to Article 34, to the contrary it further mitigates risks for the youngest cohort of users on Snapchat. As described in the Ongoing Risk Detection and Management section, Snap's DSA Independent Compliance Function and Cross-Functional Governance Team continue to regularly evaluate the effectiveness of our measures as we look to maintain or further reduce prevalence, detect any new risks, assess any deployed functionalities for critical impacts and determine whether further mitigating measures might be required.

In summary, Snap has carried out our fourth annual risk assessment of Snapchat's in-scope services required by Article 34(1) of the DSA. Snap has observed further reductions in the prevalence across our illegal and otherwise violating content categories, and our second external DSA audit concluded that Snapchat complied with the applicable DSA specified requirements during the DSA Examination Period as set out in Chapter III of the DSA, in all material respects. Snap continues to conclude that we have reasonable, proportionate and effective mitigation measures, tailored to the specific systemic risks identified, as required by Article 35(1) of the DSA.

8. Final Words

This Report records Snap's fourth Annual Risk Assessment under Article 34 of Regulation (EU) 2022/2065 (the Digital Services Act), covering the assessment cycle ending 30 June 2026. This Report is confidential and shared with the European Commission, the Dutch Digital Services Coordinator, and the relevant audit firm consistent with the applicable DSA reporting framework. Any public version is subject to applicable publication requirements and protection of confidential information.

Outlook: 2026-2027 cycle

The 2026-2027 cycle will involve both continued monitoring of Snap's systemic-risk framework and a number of regulatory and compliance developments that may inform future risk-assessment, mitigation and implementation work:

Systemic-risk monitoring priorities: operational outputs of the product and feature refinements; generative-AI and synthetic-content risk patterns; minor-protection and age-assurance developments and their effect on the Section 4.4.3 risk analysis.

Regulatory and compliance developments: further Article 28 Guidelines considerations; Directive 2024/1385/EU transposition; the Article 37 audit cycle.

Alongside these priorities and developments, the Section 4 and Section 5 framework will continue to be subject to ongoing review and refinement through the Section 6 ongoing-risk-detection processes.

Closing

Snap's DSA compliance framework is a continuing operational and regulatory commitment rather than a static end-state. The framework recorded in this Report reflects the operational reality of Snapchat's in-scope services as of the 2026 assessment cutoff, and will continue to be reviewed, calibrated and refined through the Section 6 processes and through Snap's continuing engagement with the European Commission, the Dutch Digital Services Coordinator, Member State authorities, civil-society organisations, academic research networks, and the Snapchat community.